AMERICAN
STATISTICAL ASSOCIATION

NEW SERIES, No. 112. DECEMBER, 1915.

THE SERVICE OF STATISTICS IN PROBLEMS OF WAR AND PEACE.*

By E. Dana Durand.

The great war now tearing at the vitals of the world brings vividly before us the fact that the one most important political problem is the problem of war. Is war a good or an evil? Does war if successful pay the victor? Can war in any case be escaped? What are the real causes of war? What are its real effects?

These great problems have long been the subject of much discussion. In particular during the years just preceding the present war, a great deal of literature on the subject was evoked by the enormous expenditures of nations in preparation for war and by the obvious threat of the outbreak of hostilities. Some writers, like Bernhardi, glorified war as the duty of the superior race and as a benefit not merely to such race but to the world in general. Others, like Norman Angell, urged that war is an economic, social, and moral absurdity from which no nation and no class can gain. Most writers on the subject of war and peace, however, argued primarily on *a priori* grounds or made statements of fact without adequate proof. There is profound need of more detailed and more scientific inquiry. The economists, political scientists, and statisticians of the world ought to undertake a study of problems of war and peace on a larger scale and with greater thoroughness than they have hitherto done.

The present war will furnish a new mass of data for such an investigation. From no minor war could any such evidence be derived as to the effects of warfare upon the nations con-

* Address presented at special meeting of the American Statistical Association at San Francisco, August 11, 1915.

cerned and upon the rest of the world. Moreover, it is precisely this kind of gigantic conflict which must be anticipated from time to time in the future unless the world becomes convinced that warfare is unprofitable and succeeds in devising means of checking its recurrence. There is little chance that any two first rate powers in the future can resort to arms without bringing in other powers and initiating a world conflict.

A large proportion of the facts which should be known with reference to war and which bear on war problems are statistical facts. Many of the social and economic phenomena involved can be set forth only by statistics. The collection and analysis of data bearing upon the subject constitute an enormous field for the scientific statistician. If, as nearly all people hope, the United States can keep from being drawn into the present conflict, the American statistician will be peculiarly qualified to conduct statistical investigations regarding that conflict and its results. Free from the prejudice which can hardly fail to influence even the scientists of the warring nations, he should be able to present facts in their true light.

To mention some of the specific topics within this general field which lend themselves to statistical investigation is the object of this paper.

Most obvious, of course, is the topic of the economic cost and losses of war. So far as the direct costs are concerned, no great difficulty confronts the statistical investigator. He should readily be able to ascertain the direct expenditures of governments in preparation for war, in the actual conduct of war, in paying interest on war debts and in pensions. It would seem desirable to analyze data of expenditures in very considerable detail and to present also full information as to the sources of the funds expended. There would be much interest, also, in statistics as to the quantities of the various kinds of commodities and services used for military purposes and as to the sources whence the commodities were obtained—whether from previously accumulated stocks, from home production during the war, or by importation or capture.

The indirect costs of war furnish a more difficult field of inquiry. For instance, there is the question of the cost, under the system of compulsory military training, of taking

the young men of the nation for a period of years from the farm or the factory to the camp and the barrack. It may be, as contended by some, that the value of the military training in the life of men even in times of peace more than counterbalances the loss of time. This claim seems capable of investigation in some measure by statistical methods. One might compare the efficiency of the men of nations which have compulsory military service with that of the men of other nations, taking due account of differences in other conditions.

In the second place, in reckoning the cost of actual warfare the statistician might seek to compute the money value of the loss through the destruction of human lives and the maiming and invaliding of men. This is an economic loss quite aside from the pain and sorrow entailed.

Again, the statistician might seek to calculate the losses due to the destruction of non-military property and the disturbance of production and trade through war. Of course, one must not duplicate costs by adding to the expenditures of governments for the conduct of war, the amount by which the production of commodities not used for military purposes is reduced during the conflict. What happens in time of war is that a large part of the labor and capital of the country is turned from one occupation to another. People forego the production and consumption of certain kinds of goods in order to expend an equivalent amount of energy for military ends. But this very process of turning energy from one channel to another involves enormous friction and loss both during the war period itself and for a long time thereafter, and some measure of this is doubtless possible.

Especially great is the shock to business from a world war in these modern days when so much business is international in character. Before the present war broke out a large proportion of the products of each of the warring nations was being sold in other countries, great quantities in fact in countries now become enemies. International investment of capital had become of vast importance. A complex fabric of international exchange and credit had been built up. The business interests of each country were closely interwoven with those of other lands. This intricate business relationship of nations

has been torn asunder. Dealings between hostile countries have naturally ceased. Much of the business of the belligerents with neutral countries has been broken off or radically altered in character. The economic effects of the war have extended far beyond the boundaries of the fighting nations. This tremendous upheaval will leave its traces long after the war is over. The business relationships which formerly existed will not be readily renewed; in many cases the relationships of the future will permanently differ from those of the past.

To determine as far as possible the actual effects of the present war upon international business, to measure the nature, location, and extent of the losses involved and of the gains, if any there be, is a task largely within the province of the statistician.

Still another exceedingly important subject for investigation is the effect of war upon the distribution of wealth and income. Are the costs and losses fairly distributed among the people? The distribution of the burdens of taxation for war purposes, of the burdens of military service and of the indirect costs and losses of war should be considered. The relative effects upon different economic and social classes, upon different industries and occupations, should be measured so far as possible. Among the more specific problems in this field is that of the effect upon the different classes of people of the creation of a huge national debt. This subject has already been discussed with much acumen from an abstract standpoint, but more statistical facts concerning it should be sought.

The present war will afford important material for the study of vital statistics. Its effects upon the growth of population, both in the warring and in neutral countries, upon the marriage rate and the birth rate, upon migration, are bound to be marked. Still more interesting are questions as to the influence of war upon the physical, mental, and moral characteristics of the people. It has been urged that a great war, by killing off a large proportion of the stronger men, must cause a physical deterioration in succeeding generations. It has been suggested, too, that in the same way a deterioration of mental ability and even of morals may result from the decima-

tion of the population. The statistician can do something, at least, to test the validity of these opinions. Measurements and mental tests of children born during and after the war may be compared with those of children born during times of peace. Statistics of insanity and other congenital defects should be carefully studied at this particular time. Such investigations in vital statistics may throw new light upon laws of heredity and of human reproduction. It will be recalled that some have even argued that an increase in the relative number of male compared with female children results from the destruction of a large fraction of the adult men in a great war. The present conflict will furnish a test of this very doubtful view.

Such statistical studies of the effects of war as have been suggested—and these are by no means all that might be undertaken—would have not merely scientific interest; they might materially affect the future. The direction of this influence would of course depend upon the facts ascertained. Should it appear that the costs and losses of war are greater than ordinarily supposed, that those costs and losses are very unequally distributed among the people, and that serious injury results to the vitality of the nation, a general dissemination of these facts would surely cause rulers and people to be less ready to unsheathe the sword.

In quite another way, moreover, might the statistician affect the future of war and peace. He could throw more light upon the economic causes and motives of war. We need to know more, in the first place, as to the gains which a nation entering into any policy that tends to arouse international hostility may hope to attain by such policy. In the second place, we need more information as to the effect of the progress of one nation, particularly in its foreign trade, upon the prosperity of other nations. Finally, we need more light as to the economic advantages which a nation entering into war may expect to gain if it is victorious.

These three questions are intimately related. The same statistical investigations would go far toward answering them all. To illustrate this point it may be noted that it is not necessary to await the outcome of actual conflict in

order to form some judgment as to the possible gains of victory. For the economic objects which nations aim to secure by war are usually either precisely the same as, or more or less similar to, the objects which at other times they seek by less forceful means and which not infrequently one or another nation so obtains. Thus new colonial possessions, new spheres of influence in backward lands are sometimes gotten in peaceful ways, sometimes only by the sword. If the statistician can measure the gain from those acquired in the one way, we can forecast the gain from those acquired in the other. And by much the same process we can learn whether and how much rival powers are injured by such acquisition of territory or of spheres of influence.

That statistics have an important bearing in showing the wisdom or the folly of those nationalistic policies which tend to bring about war is evident from a mere enumeration of some such policies. For example, nations often seek the partial or complete exclusion of foreign goods, or seek to give their own citizens an advantage in international trade over competitors abroad. Import and export duties, laws for the direct exclusion of foreign products, bounties direct or disguised for home production or for exportation, are some of the measures by which these objects are sought. Other nations retaliate and the mere struggle of legislation and diplomacy may develop into armed conflict, or at least become a contributing cause of it. Other policies which often become a source of international friction are those with respect to shipping and other transportation agencies, to migration and the rights of foreign residents, to the investment of capital by foreigners, and especially to colonies and spheres of influence.

Regarding the effects of all economic policies of this character, much may be learned from statistics. Data as to the production of wealth in general and of particular kinds of wealth, as to the investment of capital, as to exports and imports, as to international exchange, as to currency and banking, as to prices and wages—all of these have their bearing on the wisdom or unwisdom of such policies.

It has been contended by many that, even if such policies

as have been mentioned did not involve the risk of drawing nations into war, they would still be economically unwise from the standpoint of the nation adopting them. We are told that it cannot pay a nation to adopt any line of conduct designed to give its citizens an advantage over those of other countries. Freedom for persons, for trade and for capital, parity between the citizen and the foreigner, are held to promote the wellbeing of a given country more than any form of discrimination. Others without such sweeping condemnation of all nationalistic policies maintain that at least this, that or the other such policy must fail to achieve the advantage sought. Against these stand many firm believers in the gains from nationalistic policies skilfully pursued. Much of the discussion pro and con has consisted of mere assertions or has rested on abstract reasoning alone. The true statistician, the honest and expert analyst of mass data, has indeed devoted no little study to the problems of nationalistic policy, but he has far from exhausted the field. In some directions there is still lack of statistical ma'erial which might without undue difficulty be secured. In other cases existing statistics are too inaccurate to furnish the basis for sound conclusions. But above all, the available statistics have not been adequately analyzed with a view to their bearing upon nationalistic policy. The statistician should seek greater influence in the shaping of legislation and administrative action regarding such matters as tariff, shipping, immigration, rights of foreigners and of foreign capital, colonies and spheres of influence. Not only would he thus aid the nation in the promotion of its true interests, but in all probability the wiser action which would result would remove much of the friction between nations and lessen the likelihood of war.

The assumption often underlying nationalistic policies is that the gain of one country must mean the loss of another. The people of England, for example, very generally believe that they have been injured by the industrial progress of Germany and the rapid increase of German exports. The majority of Germans believe that England's control of vast and populous dependencies has hampered the development

of German foreign trade. International jealousies based on such beliefs are important causes of war.

This underlying assumption of nationalism has recently been brilliantly attacked by Norman Angell and others, but it can not be said to have been subjected to the rigid test of scientific analysis. As has already been suggested, those statistical investigations which would help in the wise shaping of nationalistic measures or which would check their undue extension, would incidentally serve to test this assumption. And, in the same way, such statistical investigations would affect the allurement of victory, would lessen that allurement should the facts disclosed tend to show that the possible gains are less than ordinarily supposed.

We have thus far considered only the economic problems connected with war. Can the statistician go further and throw new light on those problems which are distinctly not economic in character?

We have heard much in recent years about the relation of nationality to war. Inherent national traits and peculiarities are held largely to explain all great conflicts. We are told by some that one of the chief causes of war is failure to give due recognition to the principle of nationality in fixing the boundaries of states. On the other hand, some give to nationality an even more threatening position as the disturber of peace. To establish each nationality as a self-governing state would in their opinion by no means cause wars to cease, but might even increase their frequency. They hold that often the inborn differences and antipathies between the peoples of different states are so great that they are sure from time to time to break forth into armed conflict.

Here again we are dealing too much with theory and bald assertion. We know too little about the facts of nationality. It is largely in statistics that we must seek more knowledge. The statistician may describe and weigh more accurately than has yet been done the characteristics of the several nationalities at the present time and the differences among them. Especially can he help us to judge to what extent the physical, mental and moral peculiarities of peoples are inborn and to what extent they are due merely to environment and

to historical accident. Men often speak glibly of the traits of the Teuton, the Celt, the Slav, or the Roman as being universal to all members of the group and as being unchanging. It may be so, but the assertion lacks proof. Compare, if you can, not the Teuton living in the north with the Roman living in the south, but the Teuton and the Roman living under similar conditions of soil and climate. Compare, if you can, not the Slav living amid the traditions of Russia with the Anglo-Saxon living amid the traditions of England, but the Slav and the Anglo-Saxon living in the same country under the same traditions. The United States offers a peculiarly favorable place for making such comparisons. We have assembled here representatives of every nationality. By the study of the nationalities in America the statistician and the sociologist can contribute much to science. Professor Boaz of Columbia University, as the result of interesting statistical measurements, has reached the conclusion that important physical changes take place with comparative rapidity among immigrant peoples after their arrival in the United States. If that is true it seems likely that mental and moral changes, changes in habits and ideals often supposed to be inborn, must likewise take place. Further statistical investigations along this line are clearly desirable.

If statistical study should perchance indicate that the extent and the permanence of the differences among nationalities have been exaggerated, this conclusion, when made widely known among the masses of mankind, would surely tend somewhat to lessen international antipathies. It has been the hope of pacificists that as, with increasing education and increasing intercourse, the people of one nation came to know more about their neighbors they would find more to like in them and less to dislike and fear. Even if, on the other hand, the statistician should ascertain that great and enduring differences do exist among nationalities, yet might his work serve somewhat to lessen the chances of war by leading nations to adopt policies, both internal and external, based on recognition of those differences.

Finally, the venturesome statistician may try his mettle on the problem of the moral consequences of preparation for

war and of war itself. Statistical data in this field must always be inadequate, but they may perhaps enable us to form some approximate judgment of the relative merit of the claim of the extreme militarists who maintain that war has an ennobling influence, and that of the extreme pacificists who hold that war is an injury to all the highest interests of mankind. The present struggle will furnish more material for such investigations than has hitherto been available to the scientific student.

It may be that some of the suggestions here presented as to the service of statisticians in solving problems connected with war are Utopian. Even as to much simpler matters it is hard to get sufficient and sufficiently accurate statistics and hard to draw certain conclusions even from satisfactory data. There are many factors in these problems which are not capable of statistical presentation. But public policy is being year by year more affected by the work of the statistician. He has already no little influence in matters relating to war and it is not too much to expect that deliberate efforts to increase that influence will prove well worth while.

SOME POPULATION STATISTICS OF THE PACIFIC COAST STATES.*

By Walter F. Willcox, Ph.D., LL.D., *Cornell University.*

During the last forty years of the nineteenth century the population of the United States grew at a rate which slowly but steadily decreased. This fact is obscured by the published figures for 1870, which are undoubtedly under the truth by some hundreds of thousands. If we accept as more accurate the revised figures for that year contained in the census of 1890 and later,† the rates of increase were as follows:

1860–70	26.6
1870–80	26.0
1880–90	24.9
1890–00	20.7

During the first seventy years of our national history no such steady change in the rate of increase was manifest and none such appears in European countries. On the contrary, during the last half of the nineteenth century the rate of growth in Europe as a whole apparently increased.‡ Still the uniformity of this change in the United States for thirty or forty years had led all students of population to expect its continuance into the twentieth century. Among the few who have made a serious effort to forecast the growth of American population, for a few years or a longer period, no one, I believe, anticipated that the rate of growth between 1900 and 1910 would exceed that between 1890 and 1900. Yet that is precisely what happened. The decennial increase instead of being less than 20.7 was 21.0, or 0.3 per cent. greater than in the earlier decade. Even 21.0 is too small, because the census of 1910 referred to April 15 as the determining date, while

*This paper was presented at the meeting of the American Statistical Association held with the University of California on August 12, 1915.

† Census of 1910, *Abstract*, p. 22. Elsewhere I have given my reasons for doubting the correctness of this revision. (See Am. Econ. Assn., *Studies*, August, 1897, p. 248). I do not believe that the rate of increase during the decade including the Civil War was greater than the rate during the succeeding decade. But for the present argument that question is of little importance.

‡ "The Expansion of Europe" (p. 46) in *Studies in Philosophy and Psychology*, Houghton, Mifflin & Co. 1906.

that of 1900 referred to June 1 and, in consequence, the interval between the two censuses was only $9\frac{7}{8}$ years. When allowance for this change is made, the rate of decennial increase proves to have been 21.3, or 0.6 per cent. greater than during the preceding decade. This unexpected change in the rate of population growth tends to discredit all forecasts of our future population which assume that the increase follows any definite law.

To get some light upon this noteworthy modification in the growth of population, we may ask whether what was true of the whole country was true of the parts. For statistical purposes the country is divided into three main areas, the northern states, the southern states and the western states. The division line between north and south is Mason and Dixon's line, the Ohio, and the southern boundary of Missouri and Kansas; the division between these two and the western states is the western boundary of the Dakotas, Nebraska, Kansas, Oklahoma, and Texas. The percentage of increase of each of these three divisions for each of the decennial periods is shown in the following table:

PER CENT. INCREASE OF POPULATION.

Division.	1860-70.	1870-80.	1880-90.	1890-1900.	1900-10.
North	27.6	26.9	24.9	19.0	17.7
South	21.9	21.7	21.3	22.4	19.8
West	80.1	58.5	75.5	31.9	66.7

The preceding figures show that the rate of growth of the northern states as a whole fell steadily throughout the half century. The rate of growth of the southern states has changed irregularly and very slightly, its total fall being little more than one fifth of that in the North. Perhaps there was an undercount in the southern states in 1890 as well as in 1870; at least various bits of evidence suggest that as a possibility. The figures show also that since 1890 the southern states have grown faster than the northern. The western states have had high and very irregular rates of growth, with a minimum during the last decade of the nineteenth century and a sharp increase ten years later.

This analysis indicates that between 1900 and 1910 the rates of increase in North and South, just as might have been expected, were less rapid than in any previous decade. It shows also that the unforeseen rise in the rate of growth of the United States in those ten years was due entirely to the fact that in the later decade the West increased more than twice as fast as in the earlier. The increase in the West raised the average for the whole country between 1890 and 1900 by 0.6 per cent. and between 1900 and 1910 by 2.6 per cent. In other words, the greater population and the higher rate of increase in the West in combination gave it more than four times the influence upon the entire country between 1900 and 1910 that it had ten years earlier.

When the West is divided into its two main components, the Pacific states, Washington, Oregon and California, and the Mountain states, and the figures found for each, the rate of growth in the Pacific states proves to have been higher between 1900 and 1910 than in any of the preceding four decades and for the first time higher than that of the Mountain states.

The striking increase in the rate of growth of the country as a whole, of the western division and especially of the Pacific Coast states since 1900, is one of the most noteworthy results and perhaps the most unexpected result of the Thirteenth Census. The general reasons for it are doubtless better known to residents of these states than they can be to one who now sees the West for the first time. Most of them also are incapable of detection or measurement by statistics and with such I have little concern. The general question which I hope to elucidate with certain figures is: "Why have the Pacific Coast states grown so rapidly since 1900?" the special question to which this paper is limited is: "What light do statistics throw upon the sources and character of that increase and the probability of its continuance?"

In the world as a whole an increase of population results from an excess of births over deaths and in no other way. But in any particular area of the earth's surface, like the United States or the Pacific Coast states, such an increase may result either from an excess of births over deaths (natural increase)

or from an excess of immigrants over emigrants (migratory increase). Usually, however, it is a resultant of these two forms.

This distinction between natural increase and migratory increase raises the question, "Is the recent growth of the Pacific Coast states due to either alone or to a combination of the two and, if to a combination, then what proportion of it is due to each source?"

For the direct study of migratory increase, there are no complete records of arrivals in or departures from these states. Immigrants arriving at the Pacific ports of the United States are registered, but how long they remain in the Pacific states and by what routes they depart, if they depart at all, are unknown elements of the problem. Arrivals by sea constitute probably a small fraction of the total arrivals. Those arriving by land come in part across the northern and southern boundaries of the country, but more largely from or through the easterly states. For this branch of the inquiry, then, no significant figures exist and we are compelled to content ourselves with the meager results which can be reached indirectly from a study of the census figures of birthplace and residence, to which I shall return.

For the direct study of natural increase the conditions look more favorable. Since 1906 California has had a state system of registering births and deaths and since 1908 Washington has had a similar system of registering deaths. Oregon has not yet introduced state registration in either form, but it would seem as if indications at least might be derived from the records of the other two states. Certainly the material deserves to be probed before accepting or rejecting it.

I begin with California where the records have existed longer than in Washington and include births as well as deaths. They have shown the following results:

BIRTHS AND DEATHS IN CALIFORNIA, 1906–1914.

Year.	Births.	Deaths.	Excess of		Birth Rate.	Death Rate.
			Births.	Deaths.		
1906	20,974	29,303		8,329	10.3	14.1
1907	24,674	31,095		6,421	11.6	14.6
1908	28,077	31,287		3,210	12.7	14.1
1909	30,828	30,985		157	13.4	13.4
1910	32,138	32,398		260	13.4	13.5
1911	34,828	34,012	816		14.0	13.7
1912	39,330	36,709	2,621		15.2	14.2
1913	43,852	38,599	5,253		16.4	14.4
1914	46,012	37,537	8,475		16.7	13.4

In eight years the births recorded have more than doubled, while the deaths have increased only 28 per cent. The increase in the births must be due mainly to the growing completeness of registration. This is shown by the fact that even now the birth rate of the state is incredibly low. Probably the true birth rate in California is not below that in France or Ireland. If so, 20 births a year to each 1,000 population may safely be assumed as a minimum. But the rate resulting from the figures of 1914 is only 16.7, or five sixths of a rate of 20 per 1,000. Again, the number of children under one year of age living in the state at any one moment must be approximately equal to the number of births in the preceding year diminished by the number out of this group who died before the day of enumeration was reached. The number of infants under one year of age enumerated in California April 15, 1910, was 40,336; if we assume that this number is not more than nine tenths of the number of births during the preceding year, the true number of births in California between April 15, 1909, and April 15, 1910, was not less than 45,000. But the number recorded in 1909 was 30,828, and in 1910 32,138, an annual average of 31,500. This estimate indicates that in 1909–10 not more than seven tenths of the births in California were registered. Certainly not all the births and probably not more than five sixths of them are now registered. Washington has no significant state records of births. Hence none of the three states has trustworthy birth rate figures, a defect which they share with the majority of other American states.

In case of a death, the body must be so bestowed as to reduce the chance that legal complications will arise or the health of survivors be endangered. Hence ordinarily records of deaths are more complete than those of births. Are the death records of California and Washington helpful in our problem? The death rates for the two states since their beginning are as follows:

Year.	Death rate of California.	Washington.
1906	14.1	?
1907	14.6	?
1908	14.1	9.3
1909	13.4	9.8
1910	13.5	10.0
1911	13.7	8.9
1912	14.2	7.9
1913	14.4	8.5

It will be noticed that between 1909 and 1913 the death rate of California rose slowly but steadily. This is no evidence of an actual increase of mortality and thus of a deterioration in the conditions within the state. It is rather a change in which all citizens of California may take pride as evidence that by coöperation of state and local authorities an increasing proportion of all the deaths which occur in the state were recorded.

The death rates of California and Colorado since 1906 show an interesting contrast.

Year.	Death rate of California.	Washington.
1906	14.1	14.0
1907	14.6	15.3
1908	14.1	14.5
1909	13.4	14.2
1910	13.5	13.8
1911	13.7	12.9
1912	14.2	11.6
1913	14.5	11.5

Since 1909 the death rate of California has been rising, while since 1907 that of Colorado has been falling. It might seem as if Colorado had been growing more healthful and California, less healthful. But the probable meaning of the figures is that Colorado has been growing more careless and California more careful in securing the registration of all deaths occurring within its limits. A death rate of 11.5 for a state like Colorado is highly improbable. No doubt many deaths escape registration or the population of the state is greatly overestimated or both. In order to avoid the complication of estimating population for an intercensal period in regions where the rate of growth is subject to sudden and violent fluctuations and the estimates therefore are likely to be very wide of the truth, and to allow for accidental variations in a single year, let us confine ourselves to the death rates for the three year period 1909–11. As the difficulties in the way of complete registration are much greater in sparsely settled districts than in cities, the effort to judge from internal evidence whether the records of deaths in California during those years were complete has been confined to the rural districts excluding all incorporated places having at least 8,000 or 10,000 inhabitants. An average death rate of less than 12 in these districts is so low as to arouse suspicion that not all the deaths are registered. An average death rate of less than 10 is so low as to make the suspicion almost a certainty.

In the following table the counties of California have been divided into three classes: those having a death rate below 10 per 1,000 in the rural districts and no doubt with incomplete registration; those with rates between 10 and 12 in their rural districts and probably with incomplete registration; and those with rates above 12 and thus perhaps not intrinsically improbable.

AVERAGE DEATH RATES IN RURAL PORTIONS OF CALIFORNIA COUNTIES, 1909-11.

Under 10.		10-12.		Over 12.	
County.	Death Rate.	County.	Death Rate.	County.	Death Rate.
Mono	3.3	San Mateo	10.2	Monterey	12.0
Inyo	4.4	Fresno	10.4	Calaveras	12.3
Modoc	5.7	Marin	10.5	Colusa	12.4
Lassen	6.0	Contra Costa	10.6	Ventura	12.5
Imperial	6.7	Plumas	10.7	Placer	12.8
Glenn	7.0	Kings	10.9	Santa Cruz	13.0
Del Norte	8.2	San Luis Obispo	11.0	Los Angeles	13.2
Humboldt	8.2	Riverside	11.1	Sonoma	13.4
San Diego	8.2	San Benito	11.3	Mendocino	13.6
Merced	8.3	Solano	11.6	Tuolumne	13.6
Alpine	8.5	Orange	11.7	Shasta	13.9
Kern	8.9			Yolo	13.9
Sierra	8.9			Lake	14.2
Tulare	9.0			Amador	14.3
Mariposa	9.1			Trinity	14.5
Madera	9.2			Butte	14.8
Santa Barbara	9.6			Tehama	14.8
Sutter	9.7			Nevada	15.1
Siskiyou	9.8			Santa Clara	15.1
Stanislaus	9.8			San Bernardino	15.3
				Yuba	15.8
				San Joaquin	15.9
				Eldorado	16.3
				Sacramento	16.3
				Alameda	16.6
				Napa	25.4

In more than one half of the counties in California the published death rate for the rural districts is improbable and in many it is impossible.

In Washington the death rate in 1913 was only 8.5, nearly 2 per 1,000 below that in any other state. In 1906, South Dakota was admitted to the federal registration area and its death records accepted by the Census Bureau as approximately correct. Through 4 years it retained that position and its death rate was.

DEATH RATE IN SOUTH DAKOTA.

Year.	Ratio per 1,000.
1906	8.5
1907	9.1
1908	9.2
1909	8.5

After these 4 years the state was excluded from the registration area because of incomplete returns. The state of Washington now holds the unenviable position formerly occupied

by South Dakota. It might be wise for the Director of the Census to treat the Pacific Coast state in the same way in which South Dakota was treated six years ago. The inclusion of incorrect and misleading returns along with those which are significant tends to discredit the whole set of federal figures and to the uninformed suggests the unwarranted inference that Washington is far healthier than any of the 22 other states in the registration area.

But if the registration officials in the state of Washington were ready and eager to coöperate with the federal government, perhaps a compromise might be arranged whereby the state figures would be retained provisionally and so long as each annual death rate was higher than its predecessor, just as that of California has been for five years. In any state and especially in a sparsely settled state like Washington, the task of building up a system of death registration is long and difficult, and, if the process is really under way as proved by the results, I believe the Director of the Census should not exercise his discretionary power to exclude a state from the registration area.

The evidence, then, warrants the conclusion that the death records of Washington and the birth records, if not the death records, of California are very imperfect and as a result cannot serve as a basis for studying the recent growth of population in the Pacific states.

Thus far my conclusion has been the negative one that no significant materials exist for a direct study of the increase of population in the Pacific Coast states whether by excess of births over deaths or by excess of immigrants over emigrants or both.

In both fields, however, a substitute for registration figures, though only an inadequate one, exists in the census records of birthplace or of sex and age and an examination of these is now in order.

Persons living in the Pacific states grouped by birthplace are of three main classes, those born in foreign countries, those born in the Pacific states and those born in other parts of continental United States. To these should be added as a small residual class those whose birthplace was not stated,

those who were born in the United States but with state of birth unknown, those born abroad under the American flag, and those born in outlying districts of the United States, like Alaska, Hawaii, Guam, Samoa, and the Philippines. All these classes combined constituted in 1910 about 1 per cent. of the population of the Pacific states. In the following analysis they will be disregarded.

Consider first the foreign born, or immigrant population, as a whole. The increase of population in the Pacific states between 1900 and 1910 was more than 1,750,000, more than three times the amount and at more than double the rate of increase in the preceding ten years. During the same period the increase of foreign born was more than 400,000, twelve times the increase of foreign born in the preceding ten years. Of the total increase, 1900–1910, nearly one fourth (23.3 per cent.) was an increase of the foreign born population, while in the earlier decade only one seventeenth (5.8 per cent.) was due to this current. This does not prove that the number of foreign born had increased faster than the number of American born. On the contrary, the two classes increased at about the same rate. But in the decade between 1890 and 1900 immigration to these states from foreign countries had almost ceased. In that decade the number of foreign born increased only 6 per cent., while in each of the other four decades it increased by at least one half.

An introductory survey of the main birthplace classes in the population of the Pacific states at each of the last five censuses can be obtained from the following brief table.

PERCENTAGE DISTRIBUTION OF POPULATION OF PACIFIC COAST STATES, ACCORDING TO BIRTHPLACE.

Place of Birth.	1870.	1880.	1890.	1900.	1910.
Pacific states	32.8	38.9	36.6	42.4	35.8
Other continental United States	33.7	30.7	35.9	34.8	40.4
Outlying districts or unspecified			0.1	0.3	1.0
Foreign born	33.5	30.4	27.4	22.5	22.8
Total†	100.0	100.0	100.0	100.0	100.0

This shows that between 1870 and 1900 the proportion of foreign born in the population decreased from one third to little

more than one fifth, but during the following decade it remained substantially unchanged. The outstanding fact revealed by the table is the great increase between 1900 and 1910 in the proportion of the population of the Pacific states drawn from other parts of the country. In other words, the population of the Pacific states has grown so rapidly since 1900 because natives of other sections of the United States have been flocking westward. There has been no corresponding increase in the natives of the Pacific states. As a recently opened region fills with population the proportion of natives usually tends to rise and the proportion of immigrants to fall. In the preceding table the natives of the Pacific states in 1870 were less than one third, and in 1900 were more than two fifths of the population. By 1910 the proportion had sunk below what it was in 1880, 1890, or 1900. The foreign born fall into two groups styled by the census "white" and "colored." The term "colored" is used as a convenient generic term to include all racial elements other than the European whites. It embraces Negroes, Indians, Malays, Chinese, and Japanese. Whether this is good ethnology may fairly be questioned. But good ethnology is hardly to be reconciled with the requirement that its tests should be understood and, when necessary, applied by an army of more than 70,000 enumerators.

In the Pacific states the only numerous components of the foreign born colored are the Chinese and Japanese. In these states in 1910 there were 94,000 foreign born colored and 104,000 Chinese and Japanese. The excess of 10,000 Chinese and Japanese over all foreign born colored is due to the fact that thousands of the former claimed, some no doubt wrongly, to have been born in the United States. In the entire country there were nearly 20,000 Chinese and Japanese who claimed to have been born in the United States. The majority were probably residing in the Pacific states, where both elements have been domiciled for the longest time. The foreign born colored in the Pacific states decreased by 14,000 between 1890 and 1900 and increased by 22,000 between 1900 and 1910, a total increase of 8,000 in twenty years. During the same twenty years the foreign born whites increased by more than 400,000. As a result the foreign born colored are

a dwindling proportion of the total foreign born. In 1890, they were 16.7 per cent., in 1900, 13.2 per cent., and in 1910, 9.9 per cent. of the foreign born. Because of this relative decrease in the foreign born colored the proportion of whites in the total population of the Pacific states has been rising:

Date.	Per Cent White.
1860	87.0
1870	90.1
1880	89.5
1890	92.9
1900	94.9
1910	96.0

The tendency towards an increasing predominance of whites in the total population appears in the whole United States as well as in the Pacific states, but in the latter the increase since 1860 has been 9.0 per cent., while in the whole country it has been only 3.3 per cent., or little more than one third as great.

A rough estimate of the number of foreign born who arrived in the Pacific states between 1900 and 1910 may be reached by starting with the number of foreign born residing in the Pacific states in 1900 and subtracting the probable number who died before 1910. This number of deaths has been estimated by applying to the foreign born in the Pacific states the death rate of the foreign born whites in the registration area in 1900. The group was probably reduced between 1900 and 1910 by about 92,750 deaths, leaving, if emigration be ignored, about 451,600 survivors in 1910 of the foreign born who had been in the Pacific states in 1900. If this number be subtracted from the total foreign born in these states in 1910, it leaves 504,200 as the number of survivors in 1910 from the foreign born who had arrived in these states during the preceding decade. Assuming that they had been in these states on the average five years, the number of arrivals needed to produce this number of survivors was about 553,500. That this number exceeds by 235,000 the number of foreign born in the Pacific states who had been in the United States less than ten years may be explained by the fact that many foreign born who had

arrived in the United States before 1900 and had been gradually advancing across the continent reached the Pacific states after 1900.

Consider now the increase in immigrants coming from other sections of the United States or more accurately the increase in natives of other sections residing in the Pacific states. The total population increased 74 per cent.; the foreign born, 76 per cent.; those born and residing in the Pacific states increased only 46 per cent.; but the natives of other states residing in the Pacific states increased 105 per cent. Obviously the Pacific states since 1900 have recruited their numbers primarily from the natives of other states and least of all from their own loins. The immigration of natives from other states as roughly measured by these birthplace figures more than doubled during the decade. The heaviest contributions came from the Mississippi Valley states, Illinois, Missouri, and Iowa, but the most rapid increase in the flow between 1900 and 1910 was from the states farther west and the slowest increase from the states farther east and especially from New England. In order to define the contributory area more accurately I have prepared the accompanying table, showing the number of natives of each state east of them living in the Pacific states in 1900 and in 1910 with the per cent. of increase during the decade.

In the following table, Oklahoma, Florida, and Nevada are exceptional. The population residing in Oklahoma more than doubled and that born in Oklahoma increased 160 per cent. between 1900 and 1910. Probably, also, the average age of the natives of the state increased, as the average age of its residents certainly did, and as a rule migration is more common in adult life. The increased migration from Florida to the Pacific states is probably connected with the similarity of agricultural conditions in Florida and southern California, and the great prosperity of sub-tropical agriculture in California since 1900. This explanation is made more probable by the fact that the increase has gone mainly to California rather than to Oregon or Washington. Nevada shows a small increase because the emigration from that state in the years before 1900 had been abnormal.

NATIVES OF STATE SPECIFIED RESIDING IN PACIFIC STATES, 1900 AND 1910, WITH PER CENT. OF INCREASE.

State.	Number.		Per Cent. Increase.
	1900.	1910.	
Oklahoma	1,061	10,695	908.0
Colorado	6,994	26,429	277.9
North Dakota	3,939	13,691	247.5
Montana	4,035	13,762	241.6
Florida	665	2,102	216.1
South Dakota	5,199	16,376	215.0
New Mexico	1,198	3,669	206.3
Arizona	2,498	7,369	195.0
Minnesota	30,860	88,069	185.4
Idaho	5,245	14,941	184.9
West Virginia	3,962	11,143	181.2
Wyoming	1,367	3,590	162.6
Michigan	33,945	87,523	157.8
Nebraska	19,333	48,583	151.3
Texas	12,425	30,977	149.3
Georgia	3,705	9,174	147.6
Wisconsin	38,348	94,035	145.2
Utah	4,606	11,264	144.5
Kansas	33,238	75,238	126.4
North Carolina	5,456	12,205	123.7
Alabama	3,190	7,029	120.4
Arkansas	7,985	17,546	119.7
Mississippi	2,859	6,019	110.5
Iowa	62,534	131,064	109.6
Kentucky	17,228	35,456	105.9
U. S. outside of Pacific states	*824,372*	*1,691,505*	*105.2*
Illinois	81,838	166,396	103.3
Indiana	39,398	79,938	102 9
Louisiana	4,172	8,282	98.5
Tennessee	14,543	28,726	97.6
Pennsylvania	46,135	90,952	97.1
South Carolina	1,542	2,955	91.7
Missouri	69,160	131,907	90.7
Ohio	64,754	119,007	83.8
New Jersey	7,748	14,012	80.8
Delaware	966	1,701	76.1
Virginia	11,520	19,183	66.5
Connecticut	6,471	10,748	66.1
Rhode Island	2,577	4,071	58.0
District of Columbia	1,430	2,230	55.9
Maryland	5,632	8,770	55.7
Massachusetts	25,334	38,844	53.3
New York	83,562	127,813	52.9
New Hampshire	5,235	7,059	34.9
Vermont	9,001	12,027	33.6
Nevada	8,648	11,427	32.1
Maine	22,731	27,508	21.0

If the emigration of native Americans to the Pacific Coast states is analyzed with due regard to the populousness of the contributing states and the proportion of the natives of each state who were living in the Pacific Coast states to the total natives of the state living in the United States, it appears that the smallest contributions, less than 1 per cent. of the total natives, came from the southeastern states, New Jersey to Arizona inclusive in 1900 and New Jersey to Louisiana in 1910;

that the states contributing 1 to 3 per cent. of their natives in 1900 were northern states from Maine to South Dakota (with Utah), and in 1910 were a belt of states stretching across the country from New Hampshire to New Mexico; and that the states contributing more than 3 per cent. of their natives were the northern and western states from Iowa to Arizona (1900) or Michigan to Arizona (1910).

Consider now the increase between 1900 and 1910 in the natives of the Pacific states living in these states. This number rose by only 46.4 per cent., a ratio much less than the increase in any other of the population groups. This slow increase suggests either that the true death rate in the Pacific states is high, and this is not borne out by any evidence at hand, or that the true birth rate is abnormally low, and this is in harmony with what we know from other sources. To my mind, the best single substitute which the data allow for an American birth rate is the proportion of children under five years of age to the number of potential mothers, or women of age to bear children, an age which is roughly defined as 15 to 44 or 15 to 49, sometimes one group and sometimes the other being available. In the United States as a whole in 1910 there were 446 children under 5 to 1,000 women 15 to 49 years of age, and for half a century this proportion had steadily and rapidly decreased. If we divide the United States into the three parts, Pacific, Mountain, and rest of the United States, the proportions of children to 1,000 women in 1910 were as follows:

Division.	Children under 5 to 1,000 Women 15–49.	Per Cent. Deviation from Average.
Pacific	343	−23
Mountain	488	+10
Rest of United States	449	+ 1
Total	446	0

The proportion of children in the Pacific states is 23 per cent. below the average for the country, which may be taken as the normal, and 30 per cent. below that in the Mountain states, next east of this area. The conclusion that the Pacific states are characterized by an extremely low birth rate is an unavoid-

able one. The following figures show that the birth rate as thus measured has been steadily falling for half a century and is now less than half (44.5 per cent.) of what it was in 1860.

Date.	Children to 1,000 Women 15-49 Years of Age in Pacific Coast States.
1860	771
1870	653
1880	541
1890	424
1900	379
1910	343

On the whole, the evidence indicates that the astonishing increase of population in these states since 1900 has been mainly due to immigration in some degree from abroad, but especially from other states, and slightly due to excess of births over deaths. Increase as a permanent and normal social process usually depends primarily upon natural increase and a growth of population resulting mainly or entirely from immigration has a weaker guarantee of permanence, While the evidence in hand does not warrant a definite answer, it does point to a slackening in the growth of population of the Pacific states as likely to occur in the near future.

"ON THE HANDICAPPING OF THE FIRST BORN," A CRITICISM OF PROFESSOR PEARSON'S 1914 MEMOIR.*

By Louis I. Dublin, *Statistician*, and Harry Langman, *Statistical Bureau, Metropolitan Life Insurance Company, New York.*

In his paper "A First Study of the Statistics of Pulmonary Tuberculosis,"† 1907, Professor Pearson concluded among other things that the early born, particularly the first, are more liable to develop tuberculosis than are the later born. This finding of Pearson's was severely criticized by a number of workers, but especially by Yule and Greenwood,‡ Weinberg§ and Macaulay.|| In his more recent memoir, "On the Handicapping of the First Born," 1914, Professor Pearson elaborates this thesis and responds to the criticism directed at his first paper.

In his 1907 paper, Pearson classified the 381 patients of a tuberculosis sanatorium by order of birth. He also arranged the sibships, i. e. the group composed of the 381 tuberculous persons and their brothers and sisters, in a similar order. He then assumed that, in principle, these two distributions by order of birth should be identical, supposing that the distribution of the sibships is equivalent to the distribution by order of birth of the tuberculous population at large. He found actually that the first and second born among the patients were represented considerably above the number expected from the distribution of the sibships. He concluded, therefore, that the early born are more liable to tuberculosis.

It is this assumption of the identity of the two distributions

* Eugenics Lecture Series. X. Dulau and Company, London.

† Studies in National Deterioration, Drapers' Company Research Memoirs. Dulau and Company, London.

‡ On the Determination of the Size of Family and of the Distribution of Characters in Order of Birth from Samples taken through Members of the Sibships. Journal of the Royal Statistical Society, Vol. LXXVII, 1913–1914.

§ "Die rassenhygienische Bedeutung der Fruchtbarkeit" Arch. f. Rassen- u. Gesellschafts-Biologie, 1910 Vol. VII, p. 684. "Zur Frage der Messung der Fruchtbarkeit," *Ibid.*, 1913. Vol. X, p. 162.

|| The Supposed Inferiority of First and Second Born Members of Families—Statistical Fallacies. 1912. T. B. Macaulay, Montreal, Can.

that is the basis of much of the criticism. We shall also address ourselves to this matter below. In addition other more general criticisms have been made, especially by Macaulay, which indicate why we may expect the early born to be represented in larger proportions without involving Pearson's conclusion. They, the early born, have had a greater opportunity to pass through the so-called "danger zone" and to become patients of tuberculosis sanatoria; younger brothers and sisters may well have a pronounced tubercular diathesis, but because of their youth may not as yet have developed the disease sufficiently to need sanatorium care. There is also a tendency, even if not a very pronounced one, in tuberculous stocks for the undue representation of early borns; for the number of offspring in the family is likely to be cut down below the average through the early death of tuberculous parents, or through other circumstances which commonly break up family ties in tuberculous stock.

Pearson, in his second paper, virtually waives the above objections. He proceeds rather, at great pains, to substantiate the other assumption which was involved in his first paper, and which was criticized rather severely by Yule and Greenwood, namely, that the distribution of the sibships by order of birth is equivalent to the distribution of the tuberculous population at large. For, if this assumption is invalid, the two distributions referred to above are not comparable, and no valid conclusions can be drawn from any such comparison. In his reply, the author does not mathematically establish his conclusion as to the identity of the two distributions. He explains that he has considered the distributions of the affected in a number of cases where the distributions of the pathologic communities from which they were chosen were approximately known, and has compared the latter distributions with those obtained from the distributions of the affected by various methods, including those of his critics. On the basis of this comparison, the author has satisfied himself that the distribution of the sibships of the affected is in general the best representation of that of the pathologic community at large.

The author then extends his argument by introducing a

discussion of the mortality and "delicacy" rates of infants by order of birth. He also considers various indications of degeneracy in families by order of birth, including such traits as imbecility, epilepsy, insanity, albinism, criminality, tuberculosis, and congenital cataract. Using the same methods as have already been described for tuberculosis, he finds a marked weighting of the first born in all these defects, except in epilepsy.

After summing up his data, the author comes to the definite conclusion that "the small family is detrimental to race progress." This conclusion he makes without reservation; for such families are constituted to a greater degree by the early born who, as he maintains, are more often defective.

For the sake of the argument let us assume that the author's tables actually show what he claims for them. Then all that follows is that the early born of the community are more likely to be inferior than the later born. It does not follow, as Pearson infers, that the early born are necessarily inferior to other siblings of their own families. Nor is it true for the pathologic community at large, since the author's subjects, taken from institutions, are distributed mostly between certain ages. Thus, what may be true of patients in a tuberculosis sanatorium need by no means be true of tuberculous persons in general, who include many children, and these are rarely found in sanatoria. Especially in the face of evidence of inferiority of the last born, it is not safe to assume that superior individuals would be obtained by increasing the size of the families in question. As has already been pointed out, this inferiority of the early born may be due to an undue representation of small degenerate families; or it may be due to the higher average age of the elder born, the younger born not having as yet attained an age at which criminal or pathologic tendencies will have had opportunity for expression to a like degree; or to such circumstances as economic pressure upon older children, or to other external conditions, and not simply to the physiologic fact of being elder born.

Thus far we have assumed that Professor Pearson's method of procedure is valid. His method is based unequivocally on the assumption that the distribution according to order of

birth of the pathologic community from which his "marked" or affected subjects are obtained is identical with the distribution of the sibships of these subjects. For if that be the case he can use the distribution of the sibships of the affected as a norm in comparing with it the distribution of the affected, in the effort to show that actually the early born among his subjects preponderate beyond all expected proportions. We shall endeavor to show that, when there is no weighting according to order of birth among the individuals affected, the distribution of the affected or that of the pathologic community represented by them is not in any case comparable with that of their sibships. We propose to take the distribution of a normal population, and, supposing all members of it to be liable to some disease in equal proportions, obtain from it the distribution of the sibships of the affected by order of birth which is to be expected on the assumption made. We shall find that the distribution of the sibships is by necessity so different as to account for practically the whole difference found by Pearson. To illustrate our point, we submit the following table:

TABLE I

RECONSTRUCTION OF THE DISTRIBUTION, BY ORDER OF BIRTH, OF THE SIBSHIPS OF THE AFFECTED FROM THAT OF THE AFFECTED, OR FROM THAT OF THE GENERAL PATHOLOGIC COMMUNITY.

Sibling's Order. — Number in Family.	Industrial Classes, N. S. W.	Number of Families with 1, 2, or More in Family.	Individuals in Families of Respective Order.	Distribution of Sibships of Affected.	Per Mille Distribution of Affected.	Per Mille Distribution of Sibships of Affected.
	I	II	III	IV	V	VI
1.....	7,670	667	667	46,325	165.57	124.68
2.....	7,003	819	1,638	45,658	151.17	122.88
3.....	6,184	742	2,226	44,020	133.49	118.48
4.....	5,442	734	2,936	41,794	117.47	112.48
5.....	4,708	722	3,610	38,858	101.63	104.58
6.....	3,986	720	4,320	35,248	86.04	94.87
7.....	3,266	673	4,711	30,928	70.50	83.24
8.....	2,593	666	5,328	26,217	55.97	70.56
9.....	1,927	522	4,698	20,889	41.60	56.22
10.....	1,405	497	4,970	16,191	30.33	43.58
11.....	908	362	3,982	11,221	19.60	30.20
12.....	546	235	2,820	7,239	11.79	19.48
13.....	311	144	1,872	4,419	6.71	11.89
14.....	167	167	2,547	2,547	8.12	6.85
Over 14.....	209					
Total.....	46,325	7,670	46,325	371,554	1000	1000

In the foregoing table we give in Column I the distribution, by order of their birth, of the siblings of a community of 7,670 fertile families taken from the industrial class in New South Wales, as quoted in Pearson's paper on pulmonary tuberculosis. We suppose, also, that all of these siblings are equally liable to some disease irrespective of their order of birth. On this assumption, we shall obtain (Column VI) the distribution of the sibships of the affected which are to be expected from the distribution in Column I. From Column I we first obtain Column II, the distribution of the families in the community according to size. From Column I we have, for instance, 4,708 fifth born and 5,442 fourth born. But we must have a fourth born for every fifth born, and any excess of the former must therefore be due to families of four only, in this case 734 in Column II. In similar manner we obtain each of the other figures in Column II. Now we have, for instance, 734 families of four each. Then we have 2,936 individuals (Column III) all belonging to families of size four. Thus we obtain for Column III a distribution of the siblings of our population classified according to size of family to which they belong. Since we have assumed all these siblings equally liable to a disease, the figures of Column III must give, in their proportions, the distribution of the affected as well, according to size of family to which they belong.

We must realize now that the subjects under Professor Pearson's consideration (distributed as in Column III) are all included in the same institution. These subjects were presumably drawn at random from the pathologic community at large. We inquire now as to the probable number of families of the community we are considering that have more than one representative among the inmates of the institution. The probability of a particular member of the community being an inmate of the institution at any given time is small. The likelihood that two members of the same family are represented among our subjects is then extremely small, for, in the first place, both must be diseased; further, both must have the disease developed to the stage where they need institutional care; both must be sent to some institution; of all institutions, both must be committed to the par-

ticular one we are examining; both must be inmates of this institution at the same time, viz., that of our observation. We infer, then, that a very small number of the sibships of the affected are represented by more than one member in this particular institution. Hence we shall assume that for our purposes each of our affected comes from a distinct family, the error being negligibly small.

That being the case, Column III, besides giving the distribution (proportionately) of the affected in the institution according to size of family to which they belong, also gives the distribution of their families according to size. Let us suppose for a moment that the actual figures are magnified proportionately, and that the figures of Column III actually represent the number of the sibships of various sizes. In that case we should have, then, 46,325 families; hence 46,325 first borns. Again, there being a second born in every family of two or more, we add all the numbers of Column III from the bottom up to the second line and obtain, for the second figure of Column IV, 45,658 as the number of second borns. Similarly, we add all of Column III up to the third line and obtain 44,020 for the third figure of Column IV. In analogous manner we complete Column IV. These figures represent, then (magnified in proportion), the distribution by order of birth of the sibships of the affected. Column VI gives 1000 individuals distributed as are the numbers of Column IV. Now, since all siblings of our community were supposed equally liable to the disease, Column I, besides giving the distribution of our community, also gives the proportionate distribution of those affected. In Column V we have 1,000 individuals distributed as the numbers of Column I. We have, then, in Column V, the per mille distribution of the diseased, as well as that of the community at large, and in Column VI the per mille distribution of the sibships of the diseased.

The discrepancy between the last two columns is striking. Thus, as to first born, the two columns give 165.6 and 124.7 as the respective numbers of affected and their sibships on a per mille basis. A like difference is found for second and third borns. These differences in proportionate distribution were obtained on the assumption that siblings of all orders of

birth were equally liable to disease—that is, that there is no weighting on the early borns. It is not surprising then that Professor Pearson should obtain such differences in his own tabulations, and any conclusions based on such a discrepancy fall to the ground.

We shall further illustrate our point by applying formally the method used in obtaining the above table to the distribution of the 381 patients of a sanatorium used by Pearson in his 1907 paper. We obtain the following table, where Roman numerals denote columns similar to those in Table I:

TABLE II

RECONSTRUCTION OF THE DISTRIBUTION OF THE SIBSHIPS OF 381 TUBERCULOUS INDIVIDUALS IN A SANATORIUM FROM THAT OF THE TUBERCULOUS INDIVIDUALS THEMSELVES, AND A COMPARISON WITH THE SIBSHIP DISTRIBUTION AS ACTUALLY FOUND BY PEARSON.

Sibling's Order. —— Number in Family.	Distribution of Affected in Sanatorium.	Number of Families of Respective Sizes Represented by I.	Number of Affected Belonging to Families of Various Sizes.	Expected Distribution of the Sibships of the 381 Affected.	381 Distributed as the Sibships of the Affected.	Distribution of the Sibships of the Affected as Actually Found by Pearson.
	I and V	II	III	IV	VI	VII
1.....	113	34	34	381	69.69	67.1
2.....	79	38	76	347	63.47	64.4
3.....	41	−11	−33	271	49.57	58.5
4.....	52	13	52	304	55.60	50.9
5.....	39	21	105	252	46.09	43.5
6.....	18	0	0	147	26.89	32.6
7.....	18	9	63	147	26.89	22.2
8.....	9	6	48	84	15.36	15.1
9.....	3	0	0	36	6.58	10.0
10.....	3	0	0	36	6.58	6.2
11.....	3	2	22	36	6.58	3.7
12.....	1	0	0	14	2.56	2.6
13.....	1	0	0	14	2.56	1.6
14.....	1	1	14	14	2.56	1.1
Total....	381	113	381	2,083	381	381 (a)

(a) This total embraces several siblings of an order of birth beyond 14.

Here Column I or its equivalent, Column V, gives the distribution of the 381 subjects. Column VI gives the distribution of the sibships of these subjects that is to be expected from Column I. Column VII gives the distribution of the sibships of the affected as actually found by Pearson. The distributions of Columns VI and VII are fairly equivalent. But VI was obtained on the assumption that all siblings were equally liable to the disease. Hence the discrepancy that

Pearson found was to be expected on the assumption that there be no weighting according to order of birth. Certainly, then, Professor Pearson cannot use these distributions as a basis for the contention that the early born are especially liable to be diseased.

Again, let us reverse our procedure to see what number of affected may be expected from the actual distribution of the sibships. In Table I we obtained the distribution VI from that of I or V in a very definite way. Given the distribution VI, that of V is then definitely determined. Using Pearson's figures again, we obtain the following table:

TABLE III

RECONSTRUCTION OF THE DISTRIBUTION OF THE AFFECTED IN A SANATORIUM FROM THAT OF THE SIBSHIPS OF THOSE ACTUALLY AFFECTED, AND A COMPARISON WITH THE DISTRIBUTION OF THE ACTUALLY AFFECTED AS FOUND BY PEARSON.

Sibling's Order.	Distribution of the Sibships of the Affected as Found by Pearson.	Distribution of the Affected to be Expected from that of their Sibships.	Distribution of the 381 Affected as Actually Found by Pearson.
	IV and VI.	I and V	VIII
1	381	93.92	113
2	366	78.92	79
3	332	61.92	41
4	289	47.59	52
5	247	37.09	39
6	185	24.69	18
7	126	14.86	18
8	86	9.15	9
9	57	5.52	3
10	35	3.08	3
11	21	1.68	3
12	15	1.13	1
13	9	.63	1
14	6	.40	1
15 and over	9		
Total	2,164	381	381

Column IV (VI) gives the distribution of the 2,164 sibships of the affected as actually found by Pearson. Column I (V) gives the expected distribution of the affected themselves, and is obtained from VI by reversing the procedure in the other tables. Column VIII gives the actual distribution of Professor Pearson's subjects. Comparing the last two columns, we note that the first, fourth and fifth born predominate. Considered as a whole, however, the variations are too irregular to warrant any definite conclusion. The same chance

variation that made 52 follow 41 in Column VIII may have raised the first figure to 113.

Professor Pearson may point out that, in some instances, he has found variations beyond even what we should anticipate on our assumptions. But we have seen that there are several reasons why we should *a priori* expect a certain predominance of first born among the marked (except, of course, in cases of congenital cataract, albinism, etc., characteristics acquired at birth). Proper allowance must be made for such factors before we can accept any decided difference in our order of birth distributions as really significant.

On reviewing Professor Pearson's data in the light of the foregoing, we find that his hypothesis is anything but substantiated, and that the cry against the pioneer child—at any rate, so far as Professor Pearson's data indicate—has little, if any, justification. It would seem obligatory that the proper method of procedure to be pursued in an investigation such as that which is proposed should be to consider only families of the same size, and that then only families with completed pathologic history should be included. This would eliminate the complication of including a number of early borns of degenerate stocks with low fertility. Again, if any definite results were obtained, we should have some reason to suppose them true for members of the same sibship. Weighting of large families as such (the primary source of the differences in our distributions) would be eliminated. Inadequacies in the records due to omission of the data concerning the ultimate history of young siblings with distinct pathologic diatheses would be accounted for. Any other like pitfalls or sources of inaccuracy must be done away with, or allowed for, before one can feel any assurance that his conclusions, if he then obtain any, include an element of truth.

VITAL STATISTICS WORK IN CALIFORNIA.*

By George D. Leslie, *Statistician, California Board of Health.*

Statistics have been termed the handmaid of prophecy and vital statistics the bookkeeping of humanity. The books kept in California include births, deaths, and marriages, and yield sound predictions on the great events of life—the cradle, the grave, and the wedding altar.

The State Bureau of Vital Statistics was organized in 1905, and California immediately won from the U. S. Bureau of the Census the honor of recognition as a registration state for deaths along with Pennsylvania and three others admitted in 1906. Until the admission of Washington in 1908, California was one of three registration states west of the Mississippi River and the only such state beyond the Rocky Mountains. Moreover, there is a sharp contrast between mortality statistics for the two registration states on the Pacific coast. The death rate for Washington has ranged widely between merely 7.9 and 10.0 per 1,000 population, while the rate for California has varied slightly at higher levels only from 13.4 to 14.6. The California death rate was below that for the whole registration area in 1906 to 1911, though somewhat above it in 1912 and 1913.

Birth registration, deficient at first, has improved year after year in California, the total more than doubling between 1906 and 1914 in the steady rise from 20,974 to 46,012 and the rate advancing from 10.3 to 16.7. In the race between the Stork and the Scythe, births first surpassed deaths somewhat in 1911 but by 1914 the excess of births over deaths rose to over one fifth (22.6 per cent.). The California death toll was 29,303, or 14.4 per 1,000 population in 1906, against 37,537, or only 13.4 per 1,000 in 1914, increases appearing annually except for a small decline between 1908 and 1909 and a sharp drop between 1913 and 1914, when the death total decreased by 1,062, or 2.8 per cent. Marriages numbered 21,317, or

* Paper presented at the meeting of the American Statistical Association at Stanford University, Palo Alto, California, August 13, 1915.

10.5 per 1,000 inhabitants in 1906, against 31,902, or 11.5 per 1,000 in 1914, a decrease of 1,266, or 5.5 per cent., appearing between 1907 and 1908, while all other years showed successive gains but with increases very small and rates actually declining in both 1913 and 1914.

The registration system established in 1905 has been modified by a new statute in effect August 8, 1915, that follows even more closely the Model Law on this subject and which, though altering standard blanks and local districts somewhat, will make no particular change in statistical tabulations. However, descriptions here given of blank forms and registration districts relate to conditions under the law of 1915.

The California birth and death certificates are based upon forms recommended by the Bureau of the Census and the marriage certificates follow similar lines. Each blank contains identifying particulars such as names of persons and parents as well as the place and date of the event, in addition to statistical data proper. The birth certificate includes the following statistical items: sex of child; whether twin or other plural birth; for each parent the race, age, birthplace, and occupation (by both particular kind of work and general nature of industry); number of children born to the mother as well as number still living; and also inquiries on the use of a prophylactic for ophthalmia neonatorum. The death certificate contains for each decedent the sex, race, marital condition, age, occupation (by both particular kind of work and general nature of industry), birthplace, parental birthplaces, cause of death, and length of residence (in the registration district and in California), together with special information for hospitals on source of infection. The marriage certificate likewise includes the following items for both groom and bride: race; age; former marital condition with number of present marriage; occupation; birthplace; and parental birthplaces.

The local registrars forwarding to the central bureau each month original certificates for births and deaths in their respective districts are the health officers of the 34 cities having freeholders' charters, the clerks of the remaining 205 cities and incorporated towns, and the county recorders for the outside or unincorporated portions of the several counties. Orig-

inal certificates for all marriages are likewise sent each month to the State Registrar, but only by the recorders of the 58 counties in California. The duty of filing a certificate with the local registrar is fixed by law upon the physician or midwife attending a birth within 36 hours thereafter, upon the undertaker, in case of a death, within 72 hours or before any disposition is made of the body, and upon the minister or justice performing a marriage ceremony within three days of the event.

Some idea of the extent of tabulations in the State Bureau of Vital Statistics is given by a general summary of the facts for California in 1914, to appear in the next Biennial Report. Similar figures are shown regularly in these reports for three main divisions—Northern, Central, and Southern—and eight minor geographic divisions of the state, and are generally available likewise from data published for individual counties as well as for leading cities.

Males preponderate greatly among the inhabitants of Western States, the proportion of males to 100 females having been 125.5 for California in 1910 against merely 106.0 for the United States. Consequently, among decedents here the proportion of males is over 60.0 per cent. (61.4) as compared with only about 55.0 for the whole registration area. For California births, however, the percentage of males is only 51.7, being 51.5 among white babies against 53.6 among non-Caucasians, mainly Japanese. The percentage of male births is 51.2 for California mothers, 51.1 for other Americans, and 52.4 for foreign born white mothers.

With reference to race, the percentage of non-Caucasians is 8.1 for babies born, 5.4 for persons dying, and 4.6 for those marrying. The Japanese birth total, though only 719 in 1910, has risen rapidly to 2,874 for 1914. While births and deaths of Japanese occur at various places throughout California, the Japanese weddings take place to the extent of nineteen twentieths of all at San Francisco. Picture brides from Japan find expectant grooms assembled at this port, the marriage certificates being filed in dozens after the arrival of trans-Pacific steamships.

For the white elements of the population, the nativity is as follows in per cents.:

	California.	Other States.	Foreign Born.
Births (Mothers).....	31.0	39.3	29.7
Deaths (Decedents).....	26.5	42.0	31.5
Marriages (Brides)	36.4	43.3	20.6

Comparison of percentages by nativity for white mothers and brides indicates that native Californians and other Americans show less strongly among women bearing children than among those marrying, while, conversely, foreign born women contribute relatively more to birth rates than to marriage totals. The same contrasts appear when widowed and divorced brides are omitted and the comparison is made between all mothers and only first marriage brides. It seems, therefore, that in this state as elsewhere the fecundity of foreign born women surpasses that of the natives.

In regard to age, tabulated only for deaths, it appears that people live longer or die older in California than in the United States generally, the median age for decedents in this state being 49.6 years against the last published figure of 42.7 years for the whole registration area. In 1911 to 1914 the median age for California decedents was successively 48.8, 49.2, 49.4, and 49.6, indicating an upward movement even though slight. While the median age is somewhat less for males than for females in the entire registration area, 42.2 years against 43.4, it is considerably higher for men than for women in California, 50.1 years against 47.3.

For decedents of 15 years and over the per cent. distribution by marital condition is as follows for each sex:

	Single.	Married.	Widowed.	Divorced.	Unknown.
Males.....	31.5	45.3	14.5	1.8	6.9
Females.....	11.9	48.7	36.9	1.3	1.2

In further reference to marital condition, the status of grooms and brides may be noted in the following per cents.:

	Single.	Widowed.	Divorced.
Grooms.....	83.1	8.5	8.4
Brides.....	79.2	10.0	10.8

The per cent. of divorced brides, 10.8, marks a steady annual rise from only 7.4 in 1907. Divorced women outnumber divorced men even more than widows exceed widowers. Although widows outnumbered divorced women remarrying in the period 1906 to 1911, yet for the years 1912 to 1914 the divorcees have surpassed widows in increasing degree. However, the per cent. divorced is higher for brides born in other states than for native Californians, though lowest of all among foreign born brides.

Classification of marriages by number in order shows that there are many more weddings between bachelors and widows or divorcees than between single women and widowed or divorced men. Moreover, bachelors unite with divorcees much more than with widows, although spinsters marry widowers about as much as they wed divorced men. Yet where persons previously married wed again with others of prior matrimonial experience the mate for either sex is somewhat more often one who was widowed rather than divorced.

Analysis of mortality statistics by occupations for California cannot be presented satisfactorily here, the tabulations for this state being made for 90 specific occupations and for 12 groups of diseases. Discussion of causes of death is also necessarily omitted at this time. The annual tabulations published for the state cover each of the 189 titles of the International Classification by sex, race, nativity, and age periods, while figures are also given for geographic divisions as well as for individual counties and chartered cities on total deaths classified by 28 principal causes.

Reference must be made, however, to valuable data compiled for California on the length of residence of decedents, this information being particularly important in connection with the heavy mortality from tuberculosis in this state. The "Great White Plague" causes about 14 per cent. of all deaths in the whole state and about 17 per cent. of the deaths in Southern California. Yet over half the tuberculosis victims in Southern California had lived in the state less than ten years. In fact, the length of residence was under one year for 13.4 per cent. of these decedents and less than six months for 8.7 per cent. of them in 1914, the per cents. for 1914 being much

lower than averages for preceding years on account, perhaps, of decreased migration of tuberculosis sufferers in a year of financial stringency.

Furthermore, nearly one fourth (23.4 per cent.) of all California decedents with a residence of under one year were persons who finally succumbed here to tuberculosis contracted undoubtedly elsewhere. Hence, the very healthfulness of the California climate is a factor in swelling the general death rate or apparent mortality. People come here to save or lengthen lives surely doomed elsewhere.

THEORY OF STATISTICAL TABULATION.

By G. P. Watkins.

This brief paper deals with statistical tables in their most general aspects and is therefore labelled the "theory of tabulation." But it is a product of experience and indeed was conceived and in part written as a general introduction to directions and rules of tabulation for use in a statistical office. Hence in form it consists largely of statements of how things should be done. But the purpose and function of statistical tables are the fundamental thoughts throughout.

Nature of Tabulation. The general meaning of the word "table" appears to be an even flat surface with breadth not disproportionately small in comparison with length or, concretely, an object characterized by the possession of such a surface. The arrangement of ordinary reading matter is in a line or lines, while a statistical table presents itself as a surface.

The table thus differs from the ordinary page of letter type not merely in being composed mainly of figures, but also in being readable in two dimensions, that is, at least vertically as well as horizontally. "Reading matter" may also be a list of numbers. But the arrangement of the line (or "lines") of ordinary reading matter running back and forth on the page is not on a surface plan. A line of running print can be followed but one way. Such a line is like a string of beads, but with the type (as the beads) interrupted on the parts of the string extending from right to left and in position on the string as the line passes from left to right. The reader's eye must follow the string. A statistical table, on the other hand, can be read either down or across. It utilizes the dimensions of a surface. According to this conception, a list is not a table and a single column does not constitute a table.

A table may also sometimes be read diagonally, especially one of content and form such as to show correlation. The ages of men and of their wives, the age and the grade of school children, etc., may conveniently be compared with reference to the most frequent combinations in this way.

Matter not of a statistical character may also be put into a table when there is some advantage in reading it more than one way. Numerical data, whether statistical in character or not, are frequently best so arranged. The tabular form is used to furnish data for, and facilitate the processes of, computation, as in the familiar tables of logarithms, trigonometric functions, roots and powers, etc., and in interest tables. Here compactness of form and ease of reference are the important considerations, but these are also the reasons for being of the statistical table.

The implied division of numerical tables into two species, mathematical and statistical, suggests a question as to what is the difference between the two. The answer is that the first species contains abstract numbers, and the second, numbers that are at least relatively concrete. Statistical tables consist of numbers representing quantities or degrees of *concrete* things, qualities, or events. Hence the importance of statistical units and of their definite and constant significance. Indeed, the writer would describe statistics in general as concerned with concrete numbers and quantities and their relations. It constitutes a characteristic method or methods of dealing with such numbers, and also consists of the material appropriately so dealt with. These two aspects of the subject tend to be recognized in ordinary speech by the use of a singular verb with "statistics" in the first sense, while in the second sense the word is treated as plural. For statistics, in either sense of the word, the importance of the table is evident. This conception of statistics, it may be added, has important general bearings not involved with its incidental use in the present connection.

Tabular presentation has conspicuous advantages as regards economy of space and of time: of space, wherever the same class designation or name is to be applied to a large number of items brought together in the table in a single line or a single column; of time, on the part of those seeking information on a specific point, in that, by using line and column as guides, the specific fact sought can be found directly. These uses of the tabular form are not peculiar to numerical tables.

Tabulation, like speech, is a device for expressing ideas, and in particular for expressing them compactly and in a way to facilitate comparison and show relations. Ordinary linguistic symbols, arabic and other numerical notation (including the symbolic use of position), rulings and spacial relations, and sometimes forms special to tabular notation, are all employed for this purpose. As with language generally, the tabular presentation of facts should say as much as possible with a meaning as unmistakable as possible in as small a compass as possible. There should be no ambiguity, hence, for example, blanks should mean but one thing. Expression should be as direct as possible, hence, for example, information essential to a prompt grasping of the meaning of the table should not be put in footnotes if avoidable. Reasonable conventions regarding the use of symbols should be observed.

The table is the fundamental means of presenting statistical material and is so characteristic of the method that it may be considered the matrix of statistics. Those who first gave to "statistics" its present meaning, as distinguished from its older sense of "political" science, were opprobriously called "*Tabellenknechte.*" As early as the '40's of the nineteenth century, New York State provided for the publication of railroad reports *in tabular form.** Statistical competence may well be described as knowledge and skill in making and interpreting tables of concrete numerical data.

Uses of a Statistical Table. The stub of a statistical table is most commonly a geographical classification. For groups of such classes there will usually be sub-totals which condense the more detailed classification. But the stub may consist of the names of reporting entities, as in the case of many primary tables of corporation and financial statistics. The most important statistical data for public-service corporations are usually printed in such form by the various supervising commissions, including the Interstate Commerce Commission. But for much such data, especially for the distinctively statistical as opposed to the financial part, the company unit has little significance and compilations are made by geographical or other groups of companies. Where the facts are presented

*1907 Annual Report of the N. Y. Public Service Commission for the First District, Vol. 1, p. 452.

by reporting entities, the tabular form may serve the purpose merely of saving space, but the totals, which are of more statistical interest, are best obtained, and their composition best shown, by way of a table. If it were possible to provide the necessary space, it would of course be best always to tabulate by such return or report units, so that the person who used the primary data could make his own groupings and combinations. However, especially where the enumeration or report unit is the individual or the private family, aggregate presentation is unavoidable. Hence the stub-items of a table represent classes, rarely also composite individuals. In publishing statistics of manufacturers and other private business enterprises, the presentation of the facts for one or few companies by themselves is expressly avoided as tending to reveal the operations of individual establishments to competitors. Such procedure on the part of the U. S. Census Bureau and the various bureaus of labor statistics is undoubtedly wise administratively, though the fact that a large business corporation with stock broadly owned cannot properly withhold from the public any sort of statistical or financial data that is of general interest should be recognized and doubtless will in time be accepted in practice. But at present only quasi-public corporations appear to be dealt with statistically according to this principle.

The statistical interest of a geographical stub is, of course, not of the highest rank. The consideration determining its use is the fact that a general or primary table is in the first instance a record and repository of data. Only to a very subordinate extent is it wise to attempt to exhibit relations and significance in such a table. In a derivative (analytical or text) table the interest is of course different. But the arrangement of the items even of a geographical stub may be made to serve the purpose of explanation where, for example, the order of magnitude or of density is followed. In the New York First District Public Service Commission reports, the arrangement of lighting companies within groups determined by intercorporate relations in the order of size (amount of revenues) somewhat increases the statistical interest of the stub, since it is a step towards making the table show correlation.

It also puts first the companies in which a reader is likely to be chiefly interested, thus facilitating reference—which fact is doubtless of more practical importance than the slight aid afforded to interpretation. The order of the street-railway groups of companies in the same series of reports is in a general way that of expensiveness of line construction. These touches of correlational arrangement are suggestive of a use of tabulation which seldom affects primary tables. The correlational use, however, supposes the captions as well as the stub-items arranged according to the degree of some quality, and thus it involves cross-classification. Primary tables ought to be planned with reference to such possible use. Perhaps the presentation of such cross-classifications might well take the place of some geographical detail.

A statistical table is often merely, and always incidentally, a presentation of items going to make up a total or series of totals. The separate columns may accordingly contain things having little or no relation to each other and they may be given together merely to save space by making unnecessary the repetition of the stub. The unity of a table, however, will usually mean more than this. But it is doubtless the first or simplest purpose of a table to show this or that aggregate and how it is made up. The stub-items constitute the individual or class names for the things of which the numbers are the entries. The entries are themselves usually aggregates. But it is possible to use the tabular form for a mere tally sheet, in which case the entries represent the individual things.

In general the stub-item of a statistical table stands for a group or class of things, and the stub contains the terms of a classification. Classifications in statistics, it should be noted, must be comprehensive, hence there is usually need of an "other" or "miscellaneous" class, and commonly also of an "unknown" or "not specified" class. For the rest, all the principles conducive to right classification apply to stub and caption classifications.

It is above implied that the captions, also, as well as the stub-items, will usually constitute a classification, or perhaps more than one classification. The fact that columns commonly

add across to a total column supposes this situation. The statistical table thus becomes a mode of cross-classification.

In this more highly evolved use of the tabular form, a statistical table is essentially an arrangement of numerical data by which the data are cross-classified according to two sets of terms, those of the stub and those of the captions. The device of sub-classification is also frequently introduced in the captions and stub by way of compound captions, sub-division of stub-items, and sub-totals. The more complicated classifications usually require additional tables in series.

Instead of the terms of a classification, a time series, especially a succession of years, may be used in the stub and have much the same relation to the entries, except that column totals are then not always significant. But such a table is usually derivative.

Limitations upon Tabular Presentation. Cross-classification corresponds to what is known in algebra as combination and is covered under the topic, "Permutations and Combinations." The mathematical principle is that the number of possible different combinations of one set of things or classes of things (enumerated in the stub-items, let us say) with another set (enumerated and described in the captions) is equal to the product of the number of items in each set. This gives the number of cross-classes or entry-places in the table. There should be occasion to use most of these, or else the form of the table needs revision, or at least condensation.

The fact that cross-classification is a process of combination serves to bring out an important limitation upon the possibilities of tabular presentation. It is often desirable to show the associations or combinations of the units under three classifications or sets of cases. If the third of these classifications is merely twofold, the space required is merely double what it was before. If there are 12 rubrics under the third classification, the normal requirement is for 12 times as much space, or probably 13 times as much, since a total of the 12 classes will be desirable. If the original stub provides for 30 items and there are 10 columns, a presentation of all the possible combinations with a further series of 12 classes will require 30 x 10 x 12, or 3,600 cross-classes or entry-places.

If it is desired to show completely by tabulation the relations between nativity in 12 classes, age in 10 classes, sex in 2 classes, residence in 50 classes, and occupation in 100 classes, supposing every possible combination will require an entry-place, the number of cross-classes will be 12 x 10 x 2 x 50 x 100, or 1,200,000. If the 50 residence rubrics are made the items of the stub and 10 columns may be put on a page, that would mean 500 entry places to a page. The presentation of the facts would, therefore, require 2,400 pages. But the number of rubrics under each classification is fewer than it might be desirable to use. The above computation, moreover, does not provide for totals. Of course, much space could in practice be saved by reason of the omission of provision for impossible or infrequent combinations. Young children, for example, will not be found in occupations. However, the limitations upon what we may call *complete tabulation* are evident. The size of census volumes, even with their limitations, is thus explained.

The difficulty in question is avoided by seldom attempting complete tabulation. Some of the combinations are not important or not of special interest. The classification of those in a specific occupation by nativity, for example, is of interest for comparatively few occupations and comparatively few localities. It may often be assumed that the variation within one kind of classification in terms of another classification will be so small that a presentation of the facts for all of the first class combined will sufficiently meet ordinary statistical requirements. Detailed compilations also may often be made to serve for a number of years, provided the proportions found are representative and quite constant. The frequent necessity of resorting to such methods—the necessity in particular of using alternative classification instead of cross-classification—explains why a given statistical compilation will seldom enable one to answer all the questions for which a solution is sought. The facts are contained in the returns but they cannot all be presented.*

*Table XI of the 1911 street-railway report (in the volume on transportation statistics, Volume II of the 1911 Report of the New York Public Service Commission for the First District), dealing with Accidents, shows in Division C a classification of injuries by occasion and a separate classification by degree of seriousness, but the relations between the two classifications are not shown, that is, the classifications are

A report schedule from which tabulations are made is commonly itself in tabular form and may contain a cross-classification. Only one who has had practical experience with the problem of devising a general table or tables to contain what is most important in such returns can appreciate the difficulty of obtaining satisfactory results in a limited space. But the reader is prepared for an application of the theory of mathematical combinations to such a case. If only 50 such report schedules are to be tabulated in a way to show the individual returns and supposing the schedule has 10 stub-items and 20 captions, then in order to present *all* the facts it would be necessary to provide at least 200 columns of 50-line tabular matter. Alternative tabulation, on the other hand, which would utilize only the cross and down totals of the schedule, would require 30 columns. It is assumed, of course, that the data of each schedule are themselves aggregates and that each such aggregation has interest of its own. If only the totals for the 50 returns taken together are wanted, only as many entry-places are required as are contained on one of the schedules, that is, $20 \times 10 + 31$ (for totals), or 231 in all—which is a table of modest dimensions. Enumeration schedules, it should be noted, are not often of a character to raise this question in just this form.

Detailed classification according to geography or locality is, as has been stated, not of statistical interest in proportion to the amount of space it takes in primary statistical publications. Every locality, however, has a neighborhood interest

alternative and do not make a cross-classification. Disregarding the difference between "accidents," "killed," and "injured," under occasions—which in fact presents in part the facts regarding the seriousness of the result—there are 5 occasion rubrics (as condensed from 22 in the annual report form) and 6 seriousness rubrics. To present all the possible combinations of these would require 30 columns instead of the 11 at present required. The same facts, with a somewhat more detailed classification by occasion of injury (disregarding, however, the number of accidents) are sub-divided between passengers, employees, and others in the three parts of Division D. If, instead of the mere distinction between killed and injured in Division A of this table, the subdivision provided for all classes of injuries and a total, 4 x 11, or 44 columns, would have to be added. Six columns might be dispensed with, but should be kept as totals.

A reduction of what would otherwise be the undue length of Tables XXXIII and XXXIV in the employees and wages statistics of the 1911 report for lighting companies (Volume III of the 1911 Report of the New York Public Service Commission for the First District) is effected by using a condensed stub in which systems or groups take the place usually occupied by individual companies. This requires the preliminary tabulation of company returns to get the totals thus printed. The use of the full company stub would increase the length of Table XXXIII in the ratio of approximately 8 to 36. Instead of occupying 19 pages, it would take 86.

in facts about itself, which the Census Bureau and other statistical officers feel called upon to cater to. Statisticians, also, often appear to want a great amount of local sub-division of the primary data.* This demand is largely the result of attempting to show the degree of connection between various sorts of social conditions by way of comparative cartograms or of corresponding numerical analysis, a method which is in effect crudely correlational. This purpose could be better served by re-counts of the punched cards pertaining (say) to a given city, with reference not merely to showing the relation between certain conditions and certain localities, but to tracing actual connections (the individual or family being the unit) so far as the data compared came from the same schedules. But for comparison with data from diverse sources the cruder cartogramatic method might be necessary, which, however, would be aided by an adaptable locality classification. For such purposes the re-counting of census cards by responsible private agencies, as well as by the Census Bureau itself after the decennial rush is over, ought to be facilitated and encouraged wherever it would serve a public object. Something of this sort, rather than further local detail, is the true statistical desideratum.

One way in which the Thirteenth Census meets the problem of voluminousness resulting from geographical details is interesting in this connection. Instead of such details being furnished for all the states together, the Abstract has a supplement for each state giving the geographical detail for it alone. Thus such local details are furnished only so far as they are interesting and useful to each class of readers.

With our present-day mechanical facilities for "tabulation," the process of sub-division and cross-classification of aggregates is limited rather by the degree of significance of the results, and by the cost and awkwardness of voluminous reports, than by the time required to make the necessary sortings and counts of cards already punched. While the mathematical theory of combination is a good point of departure in planning

* Cf. Robert A. Woods, Unit Accounting in Social Work, QUARTERLY PUBLICATIONS OF THE AMERICAN STATISTICAL ASSOCIATION, March, 1913, Vol. XIII, p. 361. This paper was read at the 1912 annual meeting. Unfortunately, the discussion that followed is not printed.

tables, most combinations of the terms of diverse classifications, even if they occur, have no concrete significance.

Comprehensiveness, Comparability, and Compactness as Essentials of Good Statistical Tables. The significance of a statistical table, as of statistics generally, depends very largely upon its being comprehensive for the field it covers. Truth in its statistical aspect is representativeness. The only absolute guaranty of the representative quality of an aggregate is that it reflects all the units within its scope. According to the mathematical theory of probabilities, much less is necessary, but this theory does not take account of the selective tendency of events and of observation, for which the statistician must be continually on his guard. The point is illustrated by the well-known difference in quality between results obtained by complete enumeration and those obtained from a circular letter or questionnaire.

A table should not be composed of mere samples. It is better to make it of narrow scope but comprehensive as far as it goes, *i. e.*, within its territorial or other limits. A table, furthermore, is likely to be one of a series, which should all be on the same basis or, at least, conform sufficiently to the basis of the series so that its representative quality and the comparability of its totals are not appreciably impaired. The most surely understood uniform basis, meeting all the requirements of comparability, is the comprehensive basis. When a table falls short of the basis of its fellows, but in a way not such as to compel its omission altogether, the appropriate place to indicate what is lacking is a general note. Sometimes it may be well to have two sets of totals to a table, one on the most comprehensive basis, and one less comprehensive, but such as to supply aggregates for data that, though falling short of perfect comprehensiveness, may be of qualified value in other ways, as for example, in the computing of ratios. On the other hand, if it is desirable to present information in connection with only one of a series of tables, it is well, in order to avoid impairing the comparability of one table with the others of the series, to put the data that exceed the standard scope in brackets and not take them into the totals, thus letting them be *in* the table for purposes of

reference, but not strictly *of* it. Uniform comprehensiveness upon some definable basis is the ideal standard. Even a small per cent. impairment of comprehensiveness may mean a large decrease in tabular efficiency.

The same principle applies with reference to corresponding tables for a series of years. While it is desirable that new data be made use of, full notice of a change of basis should be given and it is often well to give figures and make comparisons on both the old and the new basis for the first year of the change. Especially in derivative tables attention to comparability is imperative, without regard to cost in the way of added complexity, etc. Ratios, for example, should usually be given on both bases where there is a change. This again is a question of representativeness, though here differences between aggregates, rather than the aggregates themselves, are under consideration. How important this question is in another of its phases is illustrated by the place commonly given to averages, *i. e.*, representative numbers, as the gist, if not the substance, of statistics.

The complement of the requirement of comprehensiveness is that of compactness. It is of the essence of a table to convey a large amount of information in a small space. Hence sparcely tenanted columns are an eyesore, and blank columns, even where the original classification may have reasonably planned to use them, should not be tolerated. Blank lines are hardly less justifiable. Classifications should be revised when the data as spread out show such waste of space. Unrepresented classes may be disposed of in the notes. Sparsely tenanted columns should be consolidated, subdivisions of entries being indicated by footnotes if desirable. A "miscellaneous" column may often be employed with reference to such residual classes. It should never include more than a small per cent. of the material of the table. But sometimes the desirability of keeping up tables on a uniform plan, *e. g.*, through a series of years, may justify continuing sparse columns till a comprehensive overhauling of the form of tables is undertaken.

The table must ordinarily be planned with reference to fitting the printed page, as single-page lengthwise, single-

page upright, twin upright, or as a series of such. Hence dimensions in terms of columns and lines must often be carefully studied before being finally fixed. The large page and the resulting unwieldy size of most statistical volumes are due to the need of space for manoeuvering the tabular matter. Often the presentation in sections of what is functionally one table becomes necessary.

General Tables and Derivatives Tables Distinguished. A table serving primarily the purpose of a repository of comprehensive statistical data is distinguished as a general table, also, with reference to its being closest to the original data, as a primary table.

Derivative tables are summaries and auxiliary ratio tables. They may usually be distinguished as text or analysis tables. But some ratio tables, or at least some ratios, are often included among general tables. Derivative tables are based upon general tables and contain matter suitable for incorporation in analysis. They may vary in form from year to year according to the exigencies of the situation and according to the points emphasized in the text. Unlike the general tables they will usually contain data and comparisons, including absolute and per cent. increases, for several years. Just as general tables serve to show in terms of absolute numbers the composition of aggregates, a derivative table frequently serves the purposes of explanation correspondingly by means of per cent. distribution. If text tables contain data taken direct from returns, these are so treated because of lack of comprehensiveness in the data, or of perennial interest in that kind of data. Explanatory and qualifying statements contained in general-table footnotes should, unless unimportant, be either repeated or referred to in footnotes, or in text immediately adjacent to the text tables.

It is the common practice of statistical bureaus to number tables serially for each report. If Roman numerals are used for the general tables, arabic numerals are used for derivative tables, or *vice versa*. The United States Census has in general employed arabic numerals for the serial numbers of general tables and roman numerals for text tables, but in the Thir-

teenth Census volumes the text-table numbers are arabic and roman numerals seem to be reserved for general tables.

No strict line can be, or need be, drawn between what should go into general and what into text tables, though the fact that ratios are logically a part of the analysis gives the analytical text, if there is any such, a strong claim upon them. Grand totals certainly go with the general tables not only as closing them up but also because of their importance as a proof check. But divisional totals serving the purpose of a summary may go in either place. Ratios, too, may come to have so thoroughly well-established a place as to be in effect a part of the data that the public will expect to find in connection with the general tables. A derivative table in a report containing the corresponding primary tables is seldom to be considered a thing by itself to the extent of requiring no reference to its sources on the part of a reader who uses it carefully.

Comparisons with previous years—or with corresponding months (or other portions) of previous years—are also strictly a part of analysis, but their significance is so direct and their meaning in general so unmistakable that some of them may well be looked for in the general tables. They are made much of especially in commercial and financial statistics. The United States Census is liberal in presenting comparisons for previous decennial years in its general tables.

General or primary tables rightly occupy the largest place in most government statistical publications. Indeed, some official statisticians feel that the preparation and presentation of the primary tables is their whole duty. But some working-over of the raw material by those directly concerned with its compilation is desirable, if for no other reason than the beneficial reaction on the original data and tables consequent upon analyzing and applying them to the solution of scientific and practical problems. Proper emphasis upon the function of such statistical publications as sources does not preclude brief suggestive analysis, in addition to the necessary descriptive and cautionary remarks.

The Rounding and Abbreviation of Numbers. The use of rounded or cut-off numbers should seldom be adopted in general or primary tables, though doubtless desirable in

derivative or interpretative tables. The practice is often recommended without reference to, or due emphasis upon, this very necessary qualification.

Even in derivative tables, the giving of a large number, for example, millions of inhabitants, to the last digit would mislead by its supposed suggestion of "spurious accuracy" only in the case of a reader who would have at least equal difficulty in understanding what the rounding of the figures meant. The notion that we should print numbers showing the digits only in so far as they are known to be accurate, or on the basis of the theory of probabilities considered to be so, is impractical to the height of absurdity. The truth of the stated population of New York City—4,766,883 in 1910—is not of a nature to imply that the figure 3 in the units place has statistical significance. The statistician knows that the last four digits are neither more nor less accurate or truthful if made to read 7,000 instead of 6,883. He does not need to be reminded that the 117 has no objective or exact meaning in such an aggregate. It is seldom necessary to indicate that large numerical aggregates are approximate as to the right-hand figures.

But there is also a positive objection to the rounding of such numbers. From the point of view of statistical administration it is important that, for example, the population of a large area be the total for all its parts down to the smallest district for which separate figures are given, some of which in the instance referred to actually have less than 117 inhabitants. Rounding an absolute number is never obligatory and should never be done in a way to deprive anyone of the possibility of completely checking the number and of using for this purpose, if for no other, the unmodified original aggregate. Primary numerical data should not be rounded.

As regards ratios, too, their mechanical computation with equal ease to a larger as to a smaller number of places makes the decision of how far they should be carried a question of conventional expectations and of economy of attention rather than anything more fundamental. This statement does not refer to (and does not apply for) slide-rule computations. The carrying out of ratios to two decimal places (or for per cent. to hundredths of one per cent.) seems to be the most

satisfactory practice for most cases, so far as fractions are desirable, though only the first place will usually be itself significant, the second serving rather to qualify the first. Where three decimal places are used, the printer, and sometimes the reader, will easily mistake the point for a comma.

But much depends on how far it is the statistician's aim to make his material popular—an end that is, of course, entirely worthy in itself. The desirability of rounded and abbreviated numbers, also of the use of few numbers, in statistical exposition is chiefly of the same nature as are the claims of stylistic elegance or of force (as a writer may prefer or the conditions require) in the use of the English language. The first duty of one presenting statistical results is to be adequate and accurate; if possible it is well for him to be also elegant, or forcible, or whatever else may be desirable, in his choice of words and of numerical expressions.

The process of rounding or cutting-off numbers is by no means simple or a matter of course. On the contrary, it requires considerable statistical technique—else totals will be found not to check with items and ratios not with the data from which they are derived. It may be noted incidentally that where it may seem desirable, as frequently in the case of estimates, to round or abbreviate both a relative number and the corresponding absolute number, one cannot do both and at the same time preserve the requisite verifiable relation between the two. This fact counts against the rounding even of estimates, though some sign of approximation is in such cases especially desirable.

Tabular Notation. The rounding and abbreviation of numbers is strictly a part of the subject of tabular notation, but so fundamental as to affect the character of the statistical table as such. The word "notation" properly refers to the relation between the signs and symbols used to convey the meaning of any part of the table and the significance arbitrarily or conventionally attaching to them. To illustrate, it would seem that the last two digits, 83, of the figure for the population of New York City in 1910, preceded as they are by five other digits having the significance of position proper to them according to the arabic numerical notation, ought, without

difficulty, to be interpreted as having a different statistical significance from the figure 83 as arrived at, for example, by a careful housewife on inventorying her pieces of silverware preparatory to putting them into safe deposit, or by a dairyman counting his stock.

The signs used in tabulation are chiefly arabic numerals and the letters of the alphabet in their various appropriate combinations. The position of such a sign may be a part of the notation. The notation of a table is the language in which its import is expressed; and that language should be as direct, concise, and unambiguous as it is possible to make it.

The technique of statistical notation has not reached a high stage of development. The writer, at any rate, feels that the tendency among statisticians to treat a table as a mere repository of numbers and to indicate in footnotes any state of facts not so represented is objectionable. The absence of a report, the failure to segregate returns, the character of an entry as estimated or as incomplete—all these are matters that can be shown by appropriate signs on the face of the table. The best policy would seem to be to make the tabular entries self-explanatory to as high a degree as possible, for the purposes of the particular tabulation, by the use of word or other non-numerical sign entries where feasible. Footnotes are thus reserved to supplement or qualify both numerical and sign entries and especially are not intended to take the place of lacking numbers. But the technique of tabular notation lies outside the scope of a discussion of the general aspects of statistical tabulation.

VITAL AND MONETARY LOSSES IN THE UNITED STATES DUE TO PREVENTABLE DEATHS.

By C. H. Forsyth, Ph. D., *Ann Arbor, Mich.*

The purpose of this paper is to set forth the results of a statistical investigation of the vital and monetary losses in this country due to the occurrence of preventable deaths; to show (a) to what extent the average length of human life is affected by the occurrence of such deaths; (b) the effect on the expectation of life, or average future life time, at any age; (c) the effect on the death rate, at any age; (d) how great a monetary loss is sustained through such a death, assuming that the life of every person has a value, or that during the most productive period of his life—say from the age of twenty to seventy years—every person contributes something annually to the wealth of the community in which he lives; and (e) the total value of these losses, estimated for each age for both males and females, for the whole country.

We have obtained the results set forth in this paper through the comparison of mortality tables which we have constructed and which are based upon the different sets of conditions and hypotheses that are to be introduced and discussed; and a familiarity with what is involved in a mortality table and the expectation of life is so essential to the complete understanding of what follows that we shall now explain briefly what these terms imply.

A mortality table is first of all a table of death rates computed for the different ages of human life. In addition it also exhibits the effects of these death rates upon an arbitrarily chosen community in the following manner. First, a large number of persons is assumed to be living at exactly the same age, usually at the age of birth; then the number of deaths for this age is ascertained by multiplying the number living at the age by the corresponding death rate; and, finally, the difference between the number living, or population, and the number of deaths gives the population at the next higher age. The populations at all the other and higher ages are

found successively in the same way. If we let l_x represent the population or the number of persons living at age x, d_x the number of deaths and q_x the death rate, the mortality table would appear as follows:

Age.	Population.	Deaths.	Death Rate.
0	l_0	d_0	q_0
1	l_1	d_1	q_1
2	l_2	d_2	q_2
x	l_x	d_x	q_x
x+1	l_{x+1}	d_{x+1}	q_{x+1}

where, as explained above, $l_{x+1}=l_x-d_x$

Such a table not only shows at what age the last survivor dies but also provides a means of computing the average future life time or expectation of life. Thus a mortality table based upon the death rates known to exist in a given community sets forth concisely and clearly the mortality conditions of that community, and the mortality conditions of any number of communities are very easily and readily compared by means of such tables.

Obviously, all those that survive from any year of age to the next have each lived one year; hence, if we add together the survivors of all ages beyond any particular age we shall obtain the total number of years lived by those at that age; and, finally, if we divide this total number of years by the population at the age considered we shall obtain the average future life time or expectation of life of persons at that age.

Since a death is just as apt to occur at one time in the year of death as another it is usually assumed that a person will, in the long run, live a half year in the year of death, and this one half of a year is added to the average future life time explained above to obtain what is called the complete expectation of life. Expressed symbolically, the complete expectation of life at age x, or

$$\mathring{e}_x = \tfrac{1}{2} + \frac{l_{x+1}+l_{x+2}+l_{x+3}\div --}{l_x}$$

Some of the general though important aspects of the subject under discussion have been investigated by means of short-cut methods. The use of mortality tables, however, has two decided advantages over the short-cut methods; its accuracy insures greater confidence in the results; and the results are given for each age, whereas those of the short-cut methods are given only for the age of birth.

In preparing this paper we constructed a mortality table (Table I) based upon the actual death rates found to exist at the present time in those sections of the United States where reasonably accurate mortality statistics are available. We then constructed another mortality table (Table II) based upon mortality conditions that would exist if preventable deaths were actually prevented. A comparison of these mortality tables forms the basis of our discussion of the vital and monetary losses due to preventable deaths.

It is practically impossible to give a rigid definition of a preventable death. Even if it were possible to know all about the various causes of death, different viewpoints would lead to different decisions as to the preventability of most deaths. Instead of trying to define a preventable death in such a way, we have made use of a set of ratios or percentages collected and arranged by Professor Irving Fisher of Yale University, who sent a list of ninety diseases to each of a group of the most prominent medical authorities in this country and asked them to designate what per cent. of the deaths due to each disease they considered preventable.

To quote from Professor Fisher:* "Since the word 'preventable' implies the hypothesis of different conditions from those which actually exist it is necessary to specify what hypothetical conditions shall be implied in the term. Doubtless tuberculosis would be over 99 per cent. preventable if we should conceive as our hypothetical conditions that every individual could live on the prairies of the west, out of doors, be provided with the best of food, most congenial of tasks, and free from overwork and worry. Needless to say, the figures in the table do not imply such Utopian conditions, nor do they imply new medical discoveries. . . . The hypothetical condition se-

* Irving Fisher, Bulletin 30, Report on National Vitality.

lected for the meaning of the term 'preventable' is contained in the following definition: a ratio of preventability is the fraction of all deaths which would be avoided if knowledge now existing among well-informed men in the medical profession were actually applied in a reasonable way and to a reasonable extent. The term 'reasonable' is of course elastic, and will be somewhat differently interpreted by different persons, but, as in law, where 'reasonable care' is often used as a proviso, it is impossible to make any more specific condition."

In regard to the collection and preparation of the information, we quote also: "The estimates of preventability . . . need special explanation. In a few cases, these estimates are based on statistical experience. The great majority of them are based on clinical experience merely, without any exact statistics. They are thus in the nature of expert guesses. The experts in all cases are physicians. . . . Those who gave to the construction of these estimates the benefit of their experience, observations and reading were especially asked above all to be conservative. In order to avoid any possibility of exaggeration of their estimates in the table their average was taken, and then the estimate entered as below the average given. When, as was true in a large proportion of cases, the different estimates agreed fairly well, the average was employed, or rather the nearest figure ending in 0, or 5 next below the average. If the individual estimates diverged widely, an estimate was used below the average, favoring the conservative estimates rather than the optimistic. Also in cases where only a few estimates were obtainable, the estimate as entered was put below the average of those given." Because of the conservative way in which the ratios were prepared we shall often find it convenient in the future to speak of the corresponding prevention of deaths as "reasonable" to distinguish it from another plan to be considered later.

The ratios of preventability are given in the following table:

TABLE I.*

SHOWING FISHER'S RATIOS OF PREVENTABILITY FOR THE DISEASES ENUMERATED IN THE MORTALITY STATISTICS OF THE UNITED STATES, TOGETHER WITH THE RELATIVE IMPORTANCE OF EACH DISEASE AS INDICATED BY THE PERCENTAGE THE NUMBER OF ITS DEATHS BEARS TO THE TOTAL NUMBER OF DEATHS.

	Causes of Death.	Prominence of Disease. Per Cent. of all Deaths.	Ratio of Preventability. Per Cent.
1	Premature birth	2.0	40
2	Congenital malformation of the heart	.55	0
3	Other congenital malformations	.3	0
4	Congenital debility	2.3	40
5	Hydrocephalus	.1	0
6	Venereal diseases	.3	70
7	Diarrhea and enteritis	7.74	60
8	Measles	.8	40
9	Acute bronchitis	1.1	30
10	Broncho-pneumonia	2.4	50
11	Whooping cough	.9	40
12	Croup	.3	75
13	Meningitis	1.6	70
14	Diseases of larynx—not laryngitis	.07	40
15	Laryngitis	.06	40
16	Diphtheria	1.4	70
17	Scarlet fever	.5	50
18	Diseases of lymphatics	.01	20
19	Tonsilitis	.05	45
20	Tetanus	.19	80
21	Tuberculosis—not of lungs	.17	75
22	Abscess	.08	60
23	Appendicitis	.7	50
24	Typhoid fever	2.0	85
25	Puerperal convulsions	.2	30
26	Puerperal septicemia	.4	85
27	Other diseases of childbirth	.36	50
28	Diseases of tubes	.1	65
29	Peritonitis	.5	55
30	Smallpox	.01	75
31	Tuberculosis of lungs	9.9	75
32	Violence	7.5	35
33	Malarial fever	.2	80
34	Septicemia	.3	40
35	Epilepsy	.29	0
36	General, ill-defined, and unknown causes (including "heart failure," "dropsy," and "convulsions")	9.2	30
37	Erysipelas	.3	60
38	Pneumonia (lobar and unqualified)	7.0	45
39	Acute nephritis	.6	30
40	Pleurisy	.27	55
41	Acute yellow atrophy of liver	.02	0
42	Obstructions of intestines	.6	25
43	Alcoholism	.4	85
44	Hemorrhage of lungs	.1	80
45	Diseases of the thyroid body	.02	10
46	Ovarian tumor	.07	0
47	Uterine tumor	.1	60
48	Rheumatism	.5	10
49	Gangrene of lungs	.03	0
50	Anemia, leukemia	.4	50
51	Chronic poisonings	.05	70
52	Congestion of lungs	.4	50
53	Ulcer of stomach	.2	50
54	Carbuncle	.03	50
55	Pericarditis	.1	10
56	Cancer of female congenital organs	.6	0
57	Dysentery	.5	80
58	Gastritis	.65	50
59	Cholera nostras	.09	50
60	Cirrhosis of liver	.9	60

* Fisher's Report, p. 104.

TABLE I—Continued.

	Cause of Death.	Prominence of Disease. Per Cent. of all Deaths.	Ratio of Preventability. Per Cent.
61	General paralysis of insane	.3	75
62	Hyatid tumors of liver	.002	75
63	Endocarditis	.8	25
64	Locomotor ataxia	.17	35
65	Diseases of veins	.04	40
66	Cancer of breast	.4	0
67	Diabetes	.8	10
68	Biliary calculi	.17	40
69	Hernia	.27	70
70	Cancer not specified	.9	0
71	Tumor	.08	0
72	Bright's disease	5.6	40
73	Embolism and thrombosis	.26	0
74	Cancer of intestines	.55	0
75	Cancer of stomach and liver	1.7	0
76	Calculi of urinary tract	.03	10
77	Cancer of mouth	.1	0
78	Heart disease	8.1	25
79	Influenza	.7	50
80	Asthma and emphysema	.23	30
81	Angina pectoris	.4	25
82	Apoplexy	4.4	35
83	Cancer of skin	.2	0
84	Chronic bronchitis	.8	30
85	Paralysis	1.0	50
86	Softening of brain	.2	0
87	Diseases of arteries	.83	10
88	Diseases of bladder	.2	45
89	Gangrene	.25	60
90	Old age	2.0	0

The area used as a basis of the investigation comprises the states Connecticut, Indiana, Maine, Massachusetts, Michigan, New Hampshire, New Jersey, New York, Rhode Island, Vermont, and the District of Columbia. These states, together with several cities in other states formed, in the years 1900–5, what is known as the "registration area." The essential and distinctive characteristic of the registration area is that its annual registration of deaths is regarded officially as having an error of less than 10 per cent. Additions have been made to this area since 1905 but it was thought best to limit the area under observation to the 11 registration states enumerated above, as these states were the only registration states to continue as such throughout the period considered—the 11 years 1900–10.

The plans for the construction of the mortality tables are merely plans for the computation of the corresponding death rates for each age. Once the death rates are computed, the rest of the mortality table is set up very easily.

The data necessary for this computation comprise the deaths for each year of age and the population at the beginning of that year.

The deaths are given directly in the government mortality statistics, but the population data given in the census reports do not pertain to the population at the beginning any more than they do at the end of the year. We usually assume the population data to refer to the population at the middle of the year and that to determine the population at the beginning of the year the ordinary population data must be increased by the number of deaths (usually one half of the total number) that have occurred since the beginning of the year.

The quotient of the deaths by the population at the beginning of the year for each age gives the death rate required. More refined methods would be necessary to compute death rates at ages in the neighborhood of the age of birth if much emphasis were to be placed upon the direct discussion of the several mortality tables, but as we are concerned here only with differences in corresponding results of two such tables, such methods are regarded as unnecessary.

From now on we shall often refer to those whose deaths are prevented as "restored" or as "restorations;" we shall refer also to those diseases which have the ratio "zero" in Table I as "unpreventable" instead of using the longer expression "diseases whose deaths are unpreventable."

The computation of the death rates of Table I which is based upon actual deaths is straightforward and involves no particular difficulty.

The plan for computing the death rates based upon mortality conditions wherein preventable deaths are prevented requires special explanation. As a concrete illustration, let us assume 1,000,000 persons to be living in a given community all at exactly the same age and that 50,000 of that number die during the succeeding year. We say that the death rate for that age is .05000. Assuming also that 5,000 of the 50,000 deaths are due, say to pneumonia, what would be the death rate for that age if 60 per cent., or 3,000 of the deaths due to pneumonia were prevented? In other words, how many deaths would now occur among the 1,000,000? Evidently,

the number of deaths would be 50,000 minus 3,000 plus whatever deaths would occur among the 3,000 restored, due to other diseases than pneumonia. The only way to determine how many deaths would occur among the restored is to assume the restored to be normal persons again and hence subject to ordinary diseases just as normal persons, in which case it is necessary merely to multiply the 3,000 by the probability of dying from the effects of diseases other than pneumonia. This probability is identically the death rate based upon the total number of actual deaths (as given by the government mortality statistics) diminished by the number of deaths due to pneumonia.

The method of computing death rates wherein the deaths from a large number of diseases are prevented, is exactly analogous to the above case of a single disease. The number of deaths that will occur among the restored is ascertained by multiplying the number of restored by the probability of dying from the effects of unpreventable diseases. We obtain this probability for each age by entering Fisher's table of ratios and ascertaining just what diseases are unpreventable at all (have the ratio "zero") and then divide the total number of deaths due to these diseases by the corresponding population.

If we let q'_x represent this probability or death rate based upon deaths due to unpreventable diseases, and r_x the number of restorations, the death rate for any age x, when preventable deaths are prevented, becomes

$$q_x = \frac{d_x - r_x + q'_x r_x}{l_x} = \frac{d_x - (1 - q'_x) r_x}{l_x} \qquad (a)$$

which, for the numerical example given above, becomes

$$q_x = \frac{50{,}000 - 3{,}000 + q'_x \cdot 3{,}000}{1{,}000{,}000} = \frac{50{,}000 - (1 - q'_x)\,3{,}000}{1{,}000{,}000}$$

We have used the death rates represented by q'_x also to construct a third mortality table (Table III) whose use and importance will be explained later.

After the three sets of death rates were computed, the three mortality tables themselves were constructed in accordance with the definition of such a table, given previously.

The data used in this investigation comprise the deaths and population by ages for the area considered. In other words, the data comprise the essential material for computing the death rates for the different ages.

The population for 1900 of the 11 registration states enumerated above is given for each age, but, as is well known' there is a great concentration at ages which are a multiple of 5, particularly those ending in 0; so the data were first combined into quinquennial age groups and then the values for each age were interpolated. The population data for 1910 were available only in quinquennial age groups.

The different processes of interpolation used are described in the section devoted to the mathematical computation of the death rates.

The population data for the years 1900 and 1910 by quinquennial age groups are as follows:

TABLE II.

POPULATION BY AGE GROUPS FOR YEARS 1900 AND 1910 FOR ELEVEN REGISTRATION STATES.

Ages.	1900.	1910.	Ages.	1900.	1910.
0	437,944	508,615	40–44	1,247,880	1,563,163
0–4	2,072,797	2,407,441	45–49	1,013,403	1,324,060
5–9	1,984,846	2,192,715	50–54	872,741	1,133,970
10–14	1,819,115	2,112,774	55–59	685,469	821,327
15–19	1,804,950	2,216,868	60–64	562,777	672,483
20–24	1,905,779	2,354,725	65–74	702,679	864,675
25–29	1,839,826	2,203,961	75–84	263,124	314,937
30–34	1,630,050	1,948,886	85–94	30,547	50,504
35–39	1,474,697	1,831,043	95–	1,919	2,163

The deaths are given for the age groups 0, 1, 2, 3, 4, 5–9, 10–19, 20–29,—"90 and above," and are given by single and total years from 1900 to 1910 inclusive as follows:

TABLE III.

DEATHS BY AGE GROUPS FOR EACH YEAR OF 1900-1910 AND THE TOTAL, FOR ELEVEN REGISTRATION STATES.

Years.	0.	1.	2.	3.	4.	5-9.	10-19.	20-29.
1900	71,117	16,866	7,439	4,691	3,416	9,242	14,169	27,546
1901	62,759	14,389	6,333	4,136	3,159	8,535	13,354	26,853
1902	62,634	14,367	6,523	3,912	2,954	8,121	12,570	25,198
1903	60,751	13,422	6,115	3,936	2,747	8,422	13,325	25,686
1904	64,805	14,213	6,242	4,030	2,982	8,712	15,009	27,198
1905	66,894	13,806	6,073	3,757	2,660	8,197	14,174	26,387
1906	70,750	15,416	6,513	3,990	2,816	7,938	13,860	26,188
1907	68,962	14,114	6,057	3,891	2,755	7,788	14,002	26,576
1908	68,881	14,003	6,093	3,821	2,653	7,546	13,144	24,958
1909	67,681	14,870	6,208	3,717	2,679	7,412	12,752	24,214
1910	72,096	15,041	6,643	4,084	2,785	7,972	13,320	25,874
Total	730,330	160,557	70,230	43,875	31,606	89,885	149,679	286,678

Years.	30-39.	40-49.	50-59.	60-69.	70-79.	80-89.	90- .
1900	28,322	27,152	30,987	37,556	38,631	21,471	3,424
1901	28,378	28,042	31,737	38,755	38,958	22,001	3,490
1902	27,075	26,707	30,236	36,789	37,030	20,147	3,247
1903	27,833	27,858	31,893	39,197	39,468	21,746	3,530
1904	29,935	29,773	33,939	42,050	42,515	23,471	3,775
1905	29,209	29,959	33,356	42,060	42,025	23,009	3,619
1906	29,923	30,639	34,173	42,485	41,712	23,286	3,633
1907	31,343	32,569	36,637	45,983	45,751	25,727	3,950
1908	28,684	30,705	35,151	43,530	43,662	24,124	3,878
1909	28,768	31,126	35,396	45,249	45,183	24,148	3,875
1910	30,014	33,211	38,539	48,067	48,672	26,178	4,223
Total	319,384	327,741	372,044	461,721	463,607	255,306	40,646

After the total deaths for each of the 145 diseases listed in the mortality statistics (by age groups) for all the area considered, and the 11 years 1900–10 were determined, these totals were multiplied by the appropriate ratios of preventability given by Professor Fisher, to obtain the number of deaths that are preventable. The following table gives the total number of preventable deaths:

TABLE IV.

TOTAL NUMBER OF PREVENTABLE DEATHS FOR THE NUMBER OF YEARS AND AREA GIVEN.

Ages.	rx.	Ages.	rx.
0	312,742	30-39	163,897
1	83,175	40-49	144,721
2	41,467	50-59	138,239
3	23,438	60-69	159,062
4	17,041	70-79	162,301
5-9	46,658	80-89	73,501
10-19	79,350	90-	9,320
20-29	160,177		

The following table gives the total number of deaths due only to unpreventable diseases (those that have the ratio "zero" in Table I:

TABLE V.

DEATHS DUE TO DISEASES REGARDED AS ABSOLUTELY UNPREVENTABLE.

Ages.	Deaths.	Ages.	Deaths.
0	32,748	30–39	15,109
1	1,396	40–49	32,111
2	702	50–59	46,415
3	411	60–69	56,755
4	339	70–79	63,779
5–9	1,071	80–89	58,794
10–19	2,902	90–	16,845
20–29	6,311		

After the populations for 1900 and 1910, as given in Table II, were averaged, and this average multiplied by the number 11 (as explained later), the final data, also the totals of Table III and the data of Tables IV and V were expressed in quinquennial age groups (as explained later), as follows:

Age.	Table II.	Age.	Table III.	Table IV.	Table V.
0	5,206,075	0	730,330	312,742	32,748
0–4	24,641,309	1	160,557	83,175	1,396
5–9	22,976,586	2	70,230	41,467	702
10–14	21,625,390	3	43,875	23,438	411
15–19	22,119,999	4	31,606	17,041	339
20–24	23,432,772	5–9	89,885	46,658	1,071
25–29	22,240,829	10–14	51,197	24,753	1,356
30–34	19,684,148	15–19	98,482	54,597	1,546
35–39	18,181,570	20–24	132,733	74,805	2,386
40–44	15,460,737	25–29	153,945	85,372	3,925
45–49	12,856,047	30–34	157,125	82,914	5,942
50–54	11,036,911	35–39	162,259	80,983	9,167
55–59	8,287,378	40–44	160,579	73,965	14,098
60–64	6,793,930	45–49	167,162	70,756	18,013
65–69	5,162,683	50–54	177,648	68,224	21,668
70–74	3,457,764	55–59	194,396	70,015	24,747
75–79	2,097,494	60–64	225,175	78,064	27,292
80–84	1,081,842	65–69	236,546	80,998	29,463
85–89	410,827	70–74	244,698	86,492	31,762
90–94	84,454	75–79	218,909	75,809	32,017
95–	22,451	80–84	166,834	53,610	30,772
		85–89	88,472	19,891	28,022
		90–	40,646	9,320	16,845

We shall now show more in detail the methods used in computing the three sets of death rates for mortality Tables 1, 2, and 3.

In all three cases we averaged the data for the 11 years in order to obtain results based upon an average year. However,

this plan of averaging the data was carried out by assuming as our average year that one whose population and whose deaths were equal in number to the corresponding totals of the 11 years.

Since the population data are issued only once in 10 years, we assumed that the population increased in arithmetical progression from the year 1900 to the year 1910. That is, we averaged the populations for 1900 and 1910, and multiplied the average by 11 to obtain the population for the 11 years. Perhaps a more satisfactory plan would be to assume that the population increased in geometrical progression, but as we are not concerned with values given directly in the mortality tables but rather with differences of such values, the extra work required to carry through such a plan was considered unnecessary.

From the data given by age groups in Tables II, III, IV, and V, the population, deaths, and restorations for each age were obtained by methods of interpolation to be explained later; then the populations at the beginning of the year for each age were found in accordance with the explanation given above.

The death rates of Table 1, which is based upon actual deaths, were then found by dividing the number of deaths (given in Table III) by the population at the beginning of the year (obtained from the data of Table II), for each age.

The death rates of Table 3 which is based upon deaths due only to unpreventable diseases, were found by dividing the number of deaths due only to unpreventable diseases (given in Table V) by the same population which was used to compute the death rates of Table 1.

The death rates of Table 2 which is based upon deaths that would occur if unnecessary deaths were prevented, were found through the use of the formula

$$q_x = \frac{d_x - (1 - q_x')r_x}{l_x} \qquad (a)$$

given above, where d_x refers to the deaths of Table III, q_x the death rates of Table 3, r_x the restorations of Table IV

and l_x the populations which were used to compute the death rates of Tables 1 and 2, for the different ages.

TABLE 1.

MORTALITY TABLE BASED UPON THE ACTUAL DEATHS OF ELEVEN REGISTRATION STATES FOR THE YEARS 1900-1910.

Age.	l_x.	d_x.	q_x.	$\mathring{e}_x$.	Age.	l_x.	d_x.	q_x.	$\mathring{e}_x$.
0	125,325	16,831	.134300	49.44	55	68,399	1,350	.019739	17.74
1	108,494	3,431	31622	56.03	56	67,049	1,443	21527	17.08
2	105,063	1,488	14165	56.84	57	65,606	1,538	23436	16.45
3	103,575	933	9001	56.64	58	64,068	1,617	25241	15.83
4	102,642	676	6585	56.15	59	62,451	1,682	26929	15.23
5	101,966	543	.005327	55.51	60	60,769	1,746	.028729	14.63
6	101,423	454	4481	54.81	61	59,023	1,807	30610	14.05
7	100,969	380	3757	54.06	62	57,216	1,864	32575	13.48
8	100,589	318	3164	53.26	63	55,352	1,923	34749	12.92
9	100,271	271	2706	52.43	64	53,429	1,987	37193	12.36
10	100,000	241	.002407	51.57	65	51,442	2,047	.039790	11.82
11	99,759	223	2244	50.69	66	49,395	2,102	42556	11.29
12	99,536	224	2246	49.80	67	47,293	2,161	45688	10.79
13	99,312	233	2342	48.91	68	45,132	2,222	49240	10.26
14	99,079	258	2605	48.03	69	42,910	2,281	53151	9.77
15	98,821	342	.003459	47.15	70	40,629	2,330	.057337	9.29
16	98,479	404	4098	46.31	71	38,299	2,339	61074	8.82
17	98,075	453	4616	45.50	72	35,960	2,397	66668	8.36
18	97,622	478	4900	44.71	73	33,563	2,415	71941	7.93
19	97,144	490	5039	43.93	74	31,148	2,424	77817	7.50
20	96,654	505	.005222	43.15	75	28,724	2,425	.084409	7.09
21	96,149	520	5406	42.37	76	26,299	2,418	91945	6.70
22	95,629	536	5608	41.60	77	23,881	2,391	10010	6.31
23	95,093	557	5860	40.83	78	21,490	2,330	10844	5.98
24	94,536	581	6144	40.07	79	19,160	2,241	11693	5.64
25	93,955	603	.006415	39.31	80	16,919	2,138	.12638	5.32
26	93,352	624	6685	38.56	81	14,781	2,021	13675	5.02
27	92,728	643	6935	37.82	82	12,760	1,888	14796	4.74
28	92,085	659	7152	37.08	83	10,872	1,741	16009	4.47
29	91,426	672	7347	36.34	84	9,131	1,580	17302	4.23
30	90,754	686	.007554	35.61	85	7,551	1,364	.18067	4.01
31	90,068	700	7772	34.88	86	6,187	1,201	19413	3.79
32	89,368	713	7980	34.15	87	4,986	1,037	20792	3.58
33	88,655	724	8165	33.42	88	3,949	879	22262	3.39
34	87,931	733	8331	32.69	89	3,070	728	23699	3.22
35	87,198	741	.008501	31.96	90	2,342	578	.24684	3.06
36	86,457	750	8672	31.23	91	1,764	461	26134	2.90
37	85,707	759	8861	30.50	92	1,303	360	27619	2.75
38	84,948	772	9086	29.77	93	943	276	29138	2.61
39	84,176	787	9350	29.03	94	668	205	30686	2.48
40	83,389	803	.009631	28.30	95	463	149	.32260	2.36
41	82,586	820	9930	27.57	96	314	106	33856	2.24
42	81,766	841	10288	26.85	97	208	74	35470	2.12
43	80,925	868	10727	26.12	98	134	50	37097	2.02
44	80,057	899	11232	25.40	99	84	33	38731	1.92
45	79,158	930	.011748	24.68	100	51	21	.40369	1.83
46	78,228	972	12421	23.97	101	30	13	42003	1.77
47	77,256	1,007	13034	23.26	102	17	7	43630	1.74
48	76,249	1,033	13543	22.56	103	10	5	45243	1.60
49	75,216	1,053	13995	21.87	104	5	2	46839	1.70
50	74,163	1,075	.014501	21.17	105	3	1	.48409	1.40
51	73,088	1,098	15020	20.47	106	2	1	49949	1.00
52	71,990	1,134	15745	19.78	107	1	1	51452	.50
53	70,856	1,190	16792	19.09	108	0			
54	69,666	1,267	18191	18.40					

We shall now explain the methods of interpolation. As all the deaths are available in the age groups 0, 1, 2, 3, 4, 5–9,

10–19, 20–29,—"90 and above," and the population in the age groups 0, 0–4, 5–9, 10–14,—60–64, 65–74,—85–94, "95 and above," columns of T_x were formed in each case by successive additions of the numbers at the different ages, such that, T_x represents the total number at ages x and above. For example, in the case of the deaths, T_{40} means the total number of deaths at ages 40 and above. This change into columns of T_x is seen to be necessary from the character of the original data.

Ordinary fifth differences were used to interpolate single values to form quinquennial age groups out of the decennial age groups.

The four values of T_x within each quinquennial age group (except, of course, the two at each end) were interpolated by what is known as Sprague's osculatory method. This method is so well explained in so many places that we deem it sufficient to state merely that its use has the advantage over ordinary differences in that not only the slope but also the curvature of the curve (representing the values considered) at the points of intersection of the different partial interpolation curves is considered. As a result, the curve representing the completed series of values becomes much less undulating in form and hence more smooth throughout.

The final leading differences for interpolating four values between u_2 and u_3 of the series of equidistant values u_0, u_1, u_2, u_3, u_4, and u_5 are as follows:

$$(1) \qquad \frac{\Delta u_0}{5} + 8\frac{\Delta^2 u_0}{5^2} + 11\frac{\Delta^3 u_0}{5^3} - 11\frac{\Delta^4 u_0}{5^4} + \frac{\Delta^5 u_0}{5^4}$$

$$(2) \qquad 1 \text{ " } + 6 \text{ " } + 1 \text{ " } + 3 \text{ "}$$

$$(3) \qquad 1 \text{ " } + 4 \text{ " } - 3 \text{ "}$$

$$(4) \qquad 1 \text{ " } - 2 \text{ "}$$

$$(5) \qquad 5 \text{ "}$$

TABLE 2.

MORTALITY TABLE SHOWING THE DEATH RATES AND EXPECTATIONS OF LIFE ASSUMING DEATHS TO BE PREVENTED ACCORDING TO THE RATIOS GIVEN IN TABLE I.

Age.	l_x.	d_x.	q_x.	$\mathring{e}_x$.	Age.	l_x.	d_x.	q_x.	$\mathring{e}_x$
0	112,537	8,686	.077188	62.11	57	81,845	1,028	.012493	21.46
1	103,851	1,583	15246	66.26	56	80,822	1,108	13714	20.72
2	102,268	594	5022	66.28	57	79,714	1,197	15020	20.00
3	101,674	426	4193	65.67	58	78,517	1,278	16272	19.30
4	101,248	307	3035	64.94	59	77,239	1,348	17453	18.61
5	100,941	249	.002459	64.13	60	75,891	1,419	.018698	17.93
6	100,692	213	2115	63.27	61	74,472	1,489	19988	17.27
7	100,479	182	1819	62.42	62	72,983	1,557	21332	16.61
8	100,297	158	1575	61.54	63	71,426	1,630	22821	15.96
9	100,139	139	1385	60.63	64	69,796	1,709	24487	15.32
10	100,000	126	.001259	59.72	65	68,087	1,788	.026255	14.69
11	99,874	118	1186	58.79	64	66,299	1,866	28148	14.07
12	99,756	116	1167	57.86	67	64,433	1,947	30222	13.47
13	99,640	120	1202	56.80	68	62,486	2,028	32462	12.87
14	99,520	129	1292	56.00	69	60,458	2,107	34855	12.29
15	99,391	160	.001610	55.07	70	58,351	2,184	.037423	11.71
16	99,231	183	1849	54.16	71	56,167	2,228	39669	11.15
17	99,048	202	2043	53.26	72	53,939	2,330	43189	10.59
18	98,846	213	2151	52.36	73	51,609	2,407	46645	10.04
19	98,633	217	2205	51.48	74	49,202	2,491	50628	9.51
20	98,416	224	.002279	50.59	75	46,711	2,574	.055109	8.99
21	98,192	231	2356	49.70	76	44,137	2,661	60294	8.49
22	97,961	240	2445	48.82	77	41,476	2,737	65981	8.00
23	97,721	250	2558	47.94	78	38,739	2,784	71872	7.53
24	97,471	262	2689	47.06	79	35,955	2,805	78005	7.07
25	97,209	274	.002817	46.18	80	33,150	2,818	.084998	6.63
26	96,935	286	2949	45.31	81	30,332	2,818	92901	6.20
27	96,649	298	3080	44.45	82	27,514	2,801	10179	5.78
28	96,351	309	3208	43.58	83	24,713	2,764	11184	5.38
29	96,042	321	3337	42.72	84	21,949	2,706	12328	4.99
30	95,721	333	.003476	41.86	85	19,243	2,538	.13192	4.62
31	95,388	346	3628	41.01	86	16,705	2,447	14650	4.25
32	95,042	359	3776	40.15	87	14,258	2,342	16427	3.89
33	94,683	370	3907	39.30	88	11,916	2,239	18792	3.56
34	94,313	380	4027	38.46	89	9,677	2,152	22235	3.27
35	93,933	390	.004154	37.61	90	7,525	1,858	.24684	3.06
36	93,543	401	4286	36.76	91	5,667	1,481	26134	2.90
37	93,142	413	4433	35.92	92	4,186	1,156	27619	2.75
38	92,729	427	4610	35.08	93	3,030	883	29138	2.61
39	92,302	445	4818	34.24	94	2,147	659	30686	2.48
40	91,857	460	.005007	33.40	95	1,488	480	.32260	2.36
41	91,397	482	5277	32.57	96	1,008	341	33856	2.24
42	90,915	507	5574	31.74	97	667	237	35470	2.13
43	90,408	533	5897	30.91	98	430	160	37097	2.03
44	89,875	561	6247	30.09	99	270	105	38731	1.94
45	89,314	592	.006633	29.28	100	165	67	.40369	1.86
46	88,722	627	7068	28.47	101	98	41	42003	1.79
47	88,095	661	7507	27.67	102	57	25	43630	1.71
48	87,434	692	7914	26.87	103	32	14	45243	1.66
49	86,742	721	8308	26.09	104	18	8	46839	1.56
50	86,021	751	.008734	25.30	105	10	5	.48409	1.40
51	85,271	782	9171	24.52	106	5	2	.49949	1.30
52	84,489	822	9728	23.74	107	3	2	51452	.83
53	83,667	876	10471	22.97	108	1	1	52912	.50
54	82,791	946	11431	22.21	109	0			

To apply this method of interpolation, the observed values of T_x are differenced, the differences divided by the appropriate power of 5 (or multiplied by the corresponding decimal)

as indicated above, and then the leading differences themselves formed in accordance with the above scheme.

The interpolated values of T_x for each interval are formed by the continued addition of the differences in the usual way, and there is a good check on the computation because the next higher quinquennial value of T_x given in the original or observed data must be reproduced at each stage.

Each quinquennial interval will have its own set of differences derived from the differences of an age ten years younger, but the easiest plan is to compute the successive fifth differences which by a continued addition to the lower differences of the preceding sub-interval (in each case) will lead to the desired interpolated values of the next interval. This plan avoids the separate computation of the leading differences for each interval to be interpolated.

The scheme of computing successive fifth differences has been suggested before but we were unable to find anywhere the formulas for computing these fifth differences, so we have derived them ourselves to be as follows:

$$4\,\Delta^5 + \Delta^6,\ -6\Delta^5,\ -6\,\Delta^5 - 6\Delta^6,\ 4\,\Delta^5 + 3\,\Delta^6,\ 5\,\Delta^5 + 5\Delta^6,$$

$$\text{where } \Delta^n = \frac{\Delta^n u_0}{5^4}.$$

After the interpolation itself is completed the subtraction of each value of T_x from the one at the age next below will give the desired column (of deaths or population) of values for each age.

As Sprague's method of interpolation requires two quinquennial periods on each side of the interval into which five sub-intervals are to be introduced, the interpolation at the ends of each set of data were performed by using ordinary third differences. For the inner of the two quinquennial periods at the ends of each set of data the interpolations were applied centrally. Thus the ordinates of the curve through the points u_0, u_1, u_2, and u_3 or

$$u_x = u_0 + x\Delta u_0 + \frac{x\,(x-1)}{2\,!}\,\Delta^2 u_0 + \frac{x\,(x-1)\,(x-2)}{3\,!}\,\Delta^3 u_0$$

for $x = 5/5,\ 6/5$—$10/5$ were differenced five times to form the

leading differences to interpolate four values between u_1 and u_2. The leading differences are as follows:

(1)	$\frac{\Delta u_0}{5} + 3\frac{\Delta^2 u_0}{5^-} - 4\frac{\Delta^3 u_0}{5^3}$
(2)	1 " $+ 1$ "
(3)	1 "
(4 and 5)	0

The ordinates of the same curve for x = 0/5—1/5, 5/5 were differenced five times also to form the leading differences for interpolating four intermediate values in the outer quinquennial periods. Thus, these interpolations are not applied centrally. The leading differences are as follows:

(1)	$\frac{\Delta u_0}{5} - 2\frac{\Delta^2 u_0}{5^2} + 6\frac{\Delta^3 u_0}{5^3}$
(2)	1 " $- 4$ "
(3)	1 "
(4 and 5)	0

The main advantage in using the above tables of leading differences, besides convenience and system, is the accompanying check upon the accuracy of the computation, just as in the application of Sprague's method.

A very little investigation will reveal the fact that the part of the mortality table at and beyond age 90 is very flexible as far as the data are concerned, but that the difference in the expectation of life at birth on two plans, even though they differ widely, is insignificant, because at the ulterior end of the table we are dealing not only with very small populations in comparison with the radix (or population at the beginning of the table), but also with the difference of two such small populations.

This fact is very important, for we tested, in various ways the possibility of lengthening the mortality tables, and concluded in every case that any significant lengthening of the table would require quite unreasonable assumptions. For

example, an assumption that the death rate shall remain constantly equal to that of age 89 after that age, lengthens the table itself scarcely a year and has no significant effect upon the expectation of life. We wish to emphasize this fact in order to make it clear that we have no choice worth while in dealing with the ulterior end of the mortality table.

TABLE 3.

MORTALITY TABLE SHOWING MORTALITY CONDITIONS UNDER THE MORE EXTREME OF ASSUMPTIONS CONSIDERED IN THIS PAPER IN REGARD TO THE PREVENTION OF DEATHS.

Age.	l_x.	d_x.	q_x.	$\mathring{e}_x$.	Age.	l_x.	d_x.	q_x.	$\mathring{e}_x$.
0	100,778	696	.006908	84.89	55	97,260	248	.002553	31.69
1	100,082	29	284	84.49	56	97,012	273	2819	30.77
2	100,053	14	144	83.51	57	96,739	298	3079	29.86
3	100,039	8	85	82.52	58	96,441	319	3307	28.95
4	100,031	8	71	81.53	59	96,122	332	3451	28.04
5	100,023	4	.000040	80.53	60	95,790	348	.003630	27.14
6	100,019	4	43	79.53	61	95,442	364	3811	26.23
7	100,015	4	47	78.54	62	95,078	383	4027	25.33
8	100,011	6	51	77.54	63	94,695	409	4314	24.43
9	100,005	5	54	76.54	64	94,286	440	4670	23.54
10	100,000	6	.000058	75.55	65	93,846	474	.005052	22.65
11	99,994	6	61	74.55	66	93,372	511	5472	21.76
12	99,988	6	64	73.56	67	92,861	552	5943	20.88
13	99,982	7	66	72.56	68	92,309	596	6457	20.00
14	99,975	7	67	71.57	69	91,713	644	7017	19.12
15	99,968	6	.000065	70.57	70	91,069	696	.007646	18.25
16	99,962	7	66	69.58	71	90,373	744	8235	17.39
17	99,955	7	68	68.58	72	89,629	819	9139	16.53
18	99,948	7	72	67.59	73	88,810	895	11080	15.68
19	99,941	8	78	66.59	74	87,915	985	11207	14.83
20	99,933	8	.000084	65.60	75	86,930	1,089	.012528	14.00
21	99,925	9	91	64.60	76	85,841	1,214	14143	13.17
22	99,916	10	100	63.61	77	84,627	1,352	15973	12.35
23	99,906	11	111	62.61	78	83,275	1,502	18038	11.54
24	99,895	12	125	61.62	79	81,773	1,666	20370	10.74
25	99,883	14	.000140	60.63	80	80,107	1,859	.023208	9.96
26	99,869	16	158	59.63	81	78,248	2,087	26666	9.18
27	99,853	18	177	58.64	82	76,161	2,354	30906	8.42
28	99,835	20	196	57.66	83	73,807	2,671	36194	7.67
29	99,815	22	217	56.67	84	71,136	3,050	42877	6.94
30	99,793	24	.000242	55.68	85	68,086	3,379	.049624	6.23
31	99,769	27	272	54.69	86	64,707	3,923	60632	5.53
32	99,742	30	303	53.71	87	60,784	4,625	76085	4.86
33	99,712	33	336	52.72	88	56,159	5,600	99715	4.21
34	99,679	37	369	51.74	89	50,559	7,061	13966	3.62
35	99,642	40	.000406	50.76	90	43,498	9,573	.22008	3.15
36	99,602	44	443	49.78	91	33,925	8,866	26134	2.90
37	99,558	49	492	48.80	92	25,059	6,921	27619	2.75
38	99,509	56	560	47.82	93	18,138	5,285	29138	2.61
39	99,453	64	643	46.85	94	12,853	3,944	30686	2.48
40	99,389	73	.000732	45.88	95	8,909	2,874	.32260	2.36
41	99,316	82	830	44.92	96	6,035	2,043	33856	2.24
42	99,234	92	927	43.95	97	3,992	1,416	35470	2.14
43	99,142	101	1018	42.99	98	2,576	956	37097	2.04
44	99,041	110	1106	42.04	99	1,620	627	38731	1.94
45	98,931	119	.001204	41.08	100	993	401	.40369	1.85
46	98,812	129	1310	40.13	101	592	249	42003	1.77
47	98,683	140	1419	39.18	102	343	150	43630	1.70
48	98,543	150	1522	38.24	103	193	87	45243	1.62
49	98,393	160	1622	37.29	104	106	50	46839	1.55
50	98,233	169	.001725	36.36	105	56	27	.48409	1.48
51	98,069	179	1826	35.42	106	29	14	49949	1.40
52	97,885	191	1949	34.48	107	15	8	51452	1.23
53	97,694	207	2116	33.55	108	7	4	52912	1.07
54	97,487	227	2324	32.62	109	3	2		.83
					110	1	1		.50

The tables were completed, however, for the purpose of computing the expectations of life at the earlier ages.

The death rates for Tables 2 and 3 are, of course, less than the corresponding rates for Table 1 below age 90, but the curves representing the rates for Tables 2 and 3 are much steeper at the higher ages than that of Table 1; hence, any effort to interpolate in the two tables beyond age 90 results in death rates greater instead of less than those found in the same way for Table 1, a result which, of course, is not tenable.

As the best estimate of the death rates at the extreme ages, we assumed the same values in Tables 2 and 3 as in Table 1. Any systematic effort to obtain values less than those thus chosen resulted in absurd values. For example, the assumption of a longer mortality table for Tables 2 and 3 than for Table 1 leads to intermediate values of the death rates in the neighborhood of age 90 and beyond, in excess of unity.

The tables of deaths and restorations afford abundant material for profitable discussion and therefore merit careful and intelligent interpretation. Most of such discussion, however important and valuable, would prove only distantly related to the purpose of this paper, and so must be passed over for the present. We shall enumerate a few of the most important facts which are closely related to the subject under consideration.

(a) A careful comparison of the number of deaths given in Table III* for each year shows that although mortality conditions on the whole are gradually improving from year to year, they are improving only because the decrease in the death rates at the ages preceding age 40, except at the age of birth, exceeds the increase at ages beyond age 40. In other words, while diseases operative at the younger ages are show-

*It should be kept in mind that the population increased about 20 per cent. from year 1900 to year 1910.

ing wonderful improvement, those operative at the more advanced ages are actually growing more destructive.

(b) Comparison of the number of restorations in Table IV with the total number of deaths of Table III indicates that most of the preventable diseases are operative principally at the earlier ages. In fact, many diseases were found, in connection with the preparation of Table IV, which showed improvement at the earlier ages and a deterioration at the more advanced ages. It is apparent that most attention has been paid by scientists to the ills and weaknesses of youth rather than those of the older ages.

(c) If we assume the ratios of preventable deaths to all deaths by ages, as indicated by Tables III and IV covering the area composed of the 11 registration states considered in this paper, to hold for the whole of the United States, we find that about 6,000,000 deaths out of a total of 14,000,000 deaths that occurred in the United States during the eleven years 1900–10, were wholly unnecessary.

It has been noted that relatively few deaths at the older ages are preventable; we believe, however, that if scientists could be made to realize fully the importance of improving mortality conditions at the higher ages, just as great an improvement is possible at these ages as is exhibited at present at the younger ages.

It is difficult to imagine what a tremendous advance would be made if death rates at the older ages could be made to decrease instead of being allowed to increase. In other words, if in estimating the annual change in mortality conditions, we were allowed to register an improvement at the older ages and hence add a measure of this improvement to the measure of improvement at the younger ages, instead of allowing the effects of an improvement at the younger ages to be canceled to a large extent by the effects of a deterioration at the older ages, the results would be marvelous.

We are compelled to deal, in this paper, wholly with losses due to preventable deaths that can be expressed statistically; perhaps the greatest loss of all is one which we have no way of measuring numerically, and that is the wealth of experience, knowledge, and general wisdom of older persons much

the greater part of which is lost at the present time through premature loss of memory, intellectual and physical weakness, and all other characteristics of premature old age.

Some of the greatest authorities on the diseases of mankind believe the *average* length of life should be from 75 to 200 years. If such should be the length of life, it is evident that much must be done that has not, as yet, even been attempted, and the present discussion should make it clear at what point the main attack should be directed—at the diseases of the older ages.

We have already stated that a mortality table expresses concisely and clearly the mortality conditions based upon circumstances wherein the corresponding death rates are known and are used to construct the given mortality table. We shall now compare the mortality tables, (a) Table 1, which is based upon deaths as they occur at present, and (b) Table 2, in which preventable deaths are assumed to have been prevented, to ascertain the vital losses that are due to the occurrence of preventable deaths.

We have brought together the death rates of Tables 1 and 2, in Table 4 for purposes of readier comparison. We have also added a column of corresponding differences of these death rates, and a column of the percentages of these differences of the death rates based on actual deaths.

If we keep in mind that the general death rate for this country is about 14 per 1,000, this fact will help us to understand more completely a discussion of the two sets of death rates.

The excess in death rates, given in Table 4, due to the occurrence of preventable deaths begins at the maximum value of 57 per 1,000 and decreases very abruptly by ages to about 3 per 1,000 at age 1 and then on to the minimum value of 1 per 1,000 at age 10; the excess then increases about 5 or 6 per 10,000 each 5 years until at age 50 the excess has accumulated to about 6 per 1,000, or almost one half the general death rate. After age 50, the excess in the death rate about doubles itself every 5 years to the end of the table.

The percentage of this excess of the existing death rate starts in at zero at the highest ages and increases gradually, as age decreases, through 30 per cent. at age 80, 35 per cent.

at age 65, 40 per cent. at age 50, to over 50 per cent. from age 35 down to age 15, being a little over 56 per cent. between ages 20 and 25. For all ages below age 15, the percentage adheres closely to 50 per cent.

TABLE 4.

COMPARISON OF THE DEATH RATES UNDER THE TWO ASSUMPTIONS THAT DEATHS ARE AND ARE NOT PREVENTED ACCORDING TO THE RATIOS GIVEN IN TABLE I.

Age.	Deaths.			Per Cent.	Age.	Deaths.			Per Cent.
	Prevented.	Not Prevented.	Difference.			Prevented.	Not Prevented.	Difference.	
(1)	(2)	(3)	(4)	(4)÷(3)	(1)	(2)	(3)	(4)	(4)÷(3)
0	.077188	.134300	.057112	.426	45	.006633	.011748	.005115	.437
1	15246	31622	16376	518	46	7068	12421	5353	432
2	5022	14165	9143	644	47	7507	13034	5527	422
3	4193	9001	4808	533	48	7914	13543	5629	417
4	3035	6585	3550	539	49	8308	13995	5687	409
5	.002459	.005327	.002868	.538	50	.008734	.014501	.005767	.398
6	2115	4481	2366	528	51	9171	15020	5849	387
7	1819	3757	1938	515	52	9728	15745	6017	383
8	1575	3164	1589	503	53	10471	16792	6321	376
9	1385	2706	1321	487	54	11431	18191	6760	371
10	.001259	.002407	.001148	.476	55	.012493	.019739	.007246	.368
11	1186	2244	1058	472	56	13714	21527	7813	363
12	1167	2246	1079	480	57	15020	23436	8416	360
13	1202	2342	1140	471	58	16272	25241	8969	356
14	1292	2605	1313	503	59	17453	26929	9476	352
15	.001610	.003459	.001849	.534	60	.018698	.028729	.010031	.350
16	1849	4098	2249	550	61	19988	30610	10622	347
17	2043	4616	2573	557	62	21332	32575	11243	345
18	2151	4900	2749	562	63	22821	34749	11928	344
19	2205	5039	2834	562	64	24487	37193	12706	342
20	.002279	.005222	.002943	.564	65	.026255	.039790	.013535	.340
21	2356	5406	3050	564	66	28148	42556	14408	338
22	2445	5608	3163	564	67	30222	45688	15466	338
23	2558	5860	3302	563	68	32462	49240	16778	341
24	2689	6144	3455	563	69	34855	53151	18296	344
25	.002817	.006415	.003598	.560	70	.037423	.057337	.019914	.347
26	2949	6685	3736	558	71	39669	61074	21405	350
27	3080	6935	3855	556	72	43189	66668	23479	352
28	3208	7152	3944	552	73	46645	71941	25296	352
29	3337	7347	4010	546	74	50628	77817	27189	350
30	.003476	.007554	.004078	.540	75	.055109	.084409	.029300	.347
31	3628	7772	4144	533	76	60294	91945	31651	344
32	3776	7980	4204	527	77	65981	10010	34119	341
33	3907	8165	4258	521	78	71872	10844	36568	339
34	4027	8331	4304	517	79	78005	11693	38925	333
35	.004154	.008501	.004347	.511	80	.084998	.12638	.041382	.328
36	4286	8672	4386	506	81	92901	13675	43849	320
37	4433	8861	4428	500	82	10179	14796	4617	312
38	4610	9086	4476	492	83	11184	16009	4825	300
39	4818	9350	4532	485	84	12328	1730	4974	287
40	.005007	.009631	.004624	.480	85	.13192	.180627	.04875	.269
41	5277	9930	4653	469	86	14650	19413	4763	246
42	5574	10288	4714	458	87	16427	20792	4365	210
43	5897	10727	4830	447	88	18792	22262	3470	156
44	6247	11232	4985	445	89	22235	23699	1464	62
					90	.24684	.24684	.000000	.000

The excess in death rates is the greatest at the age of birth, being over 57 per 1,000, or over four times the general death rate. The importance of the excess at this age is all the more enhanced by the fact that the population at this age is greater than for any other age in any community. Hence, the absolute loss in deaths at the age of birth is much greater than at any other age.

In examining the value of this excess in death rate at any age, the relative amount of population usually found at that age should be considered. For example, the excess in death rates at age 80 and beyond is 40 per 1,000 or greater, but the absolute loss measured in number of deaths is no wise as serious as this excess indicates, because the population at this age is always very small—a little over 6 per cent. of that at the age of birth in the United States.

We have constructed Table 5 giving approximate values of the losses in deaths for representative ages, together with corresponding approximate values of the excess in death rates, taken from Table 4. The populations for each age used to compute the losses in deaths are approximate values of those given in Table II. A comparison of these losses in deaths gives a good idea of the relative importance of the excess in death rates at each age.

TABLE 5.

Age.	Population 1910.	Loss in Death Rates, per 1,000.	Loss in Deaths.
0	509,000	57	29,013 (=509.57)
1	481,000	16	7,696
10	422,000	1	422
20	471,000	3	1,413
30	390,000	4	1,560
40	313,000	4½	1,409
50	227,000	5½	1,249
60	124,000	10	1,240
70	86,000	20	1,720
80	30,000	41	1,230

With the exception of the first ten years the absolute losses in deaths in Table 5 are remarkably alike in value for each age; the losses for the first ten years, however, begin with a loss in deaths at the age of birth of almost 25 times the average loss at any following age, and decrease through a relatively large loss at age 1, to the minimum value at age 10, which is only about one fourth that of any succeeding age.

We have brought together also the expectations of life of Tables 1 and 2 in Table 6 for purposes of comparison and discussion. A third column gives the corresponding differences in these expectations, expressed in years and days.

TABLE 6.

COMPLETE EXPECTATIONS OF LIFE AS BASED UPON THE TWO ASSUMPTIONS THAT DEATHS ARE AND ARE NOT PREVENTED ACCORDING TO THE RATIOS GIVEN IN TABLE I.

Age.	Deaths		Loss in		Age.	Deaths		Loss in	
	Not Prevented.	Prevented.	Years.	Days.		Not Prevented.	Prevented.	Years.	Days.
0	49.44	62.11	12	245	45	24.68	29.28	4	219
1	56.03	66.26	10	84	46	23.97	28.47	4	183
2	56.84	66.28	9	161	47	23.26	27.67	4	150
3	56.64	65.67	9	11	48	22.56	26.87	4	113
4	56.15	64.94	8	288	49	21.87	26.09	4	80
5	55.51	64.13	8	226	50	21.17	25.30	4	47
6	54.81	63.27	8	168	51	20.47	24.52	4	18
7	54.06	62.42	8	131	52	19.78	23.74	3	350
8	53.26	61.54	8	102	53	19.09	22.97	3	321
9	52.43	60.63	8	73	54	18.40	22.21	3	296
10	51.57	59.72	8	55	55	17.74	21.46	3	263
11	50.69	58.79	8	37	56	17.08	20.72	3	234
12	49.80	57.86	8	22	57	16.45	20.00	3	201
13	48.91	56.80	7	321	58	15.83	19.30	3	193
14	48.03	56.00	7	354	59	15.23	18.61	3	139
15	47.15	55.07	7	336	60	14.63	17.93	3	110
16	46.31	54.16	7	310	61	14.05	17.27	3	80
17	45.50	53.26	7	277	62	13.48	16.61	3	47
18	44.71	52.36	7	237	63	12.92	15.96	3	15
19	43.93	51.48	7	201	64	12.36	15.32	2	350
20	43.15	50.59	7	161	65	11.82	14.69	2	318
21	42.37	49.70	7	120	66	11.29	14.07	2	285
22	41.60	48.82	7	80	67	10.77	13.47	2	256
23	40.83	47.94	7	40	68	10.26	12.87	2	223
24	40.07	47.06	6	261	69	9.77	12.29	2	190
25	39.31	46.18	6	318	70	9.29	11.71	2	153
26	38.56	45.31	6	274	71	8.82	11.15	2	120
27	37.82	44.45	6	230	72	8.36	10.59	2	84
28	37.08	43.58	6	183	73	7.93	10.04	2	40
29	36.34	42.72	6	139	74	7.50	9.51	2	4
30	35.61	41.86	6	91	75	7.09	8.99	1	329
31	34.88	41.01	6	47	76	6.70	8.49	1	288
32	34.15	40.15	6	0	77	6.31	8.00	1	252
33	33.42	39.30	5	321	78	5.98	7.53	1	201
34	32.69	38.46	5	281	79	5.64	7.07	1	157
35	31.96	37.61	5	237	80	5.32	6.63	1	113
36	31.23	36.76	5	193	81	5.02	6.20	1	66
37	30.50	35.92	5	153	82	4.74	5.78	1	15
38	29.77	35.08	5	113	83	4.47	5.38		332
39	29.03	34.24	5	77	84	4.23	4.99		277
40	28.30	33.40	5	37	85	4.01	4.62		223
41	27.57	32.57	5	0	86	3.79	4.25		168
42	26.85	31.74	4	325	87	3.58	3.89		113
43	26.12	30.91	4	288	88	3.39	3.56		62
44	25.40	30.09	4	252	89	3.22	3.27		18
					90	3.06	3.06		0

According to Table 6, the expectation of life at age 10 is, at present, 51.57 years; the expectation of life at age 10 would be 59.72 years, or a little over 8 years more than it is at present, if unnecessary deaths were prevented.

The greatest difference, however, occurs at the age of birth where it is almost 13 years. In other words, the average length of life would be 62 years or 13 years longer than it is now if it were not for the occurrence of unnecessary deaths.

Professor Fisher obtains between 14.02 and 16.2 years by means of his short-cut method; hence, we are inclined to believe that his method gives somewhat exaggerated results as a rule. We have confirmed this belief in our own mind by other examples which, however, are in no other way related to the subject of this paper and are, therefore, not given here.

We have no criticism of Professor Fisher's results themselves when we consider the conservatism used in the preparation of his table of ratios. In fact, we shall point out later that this loss in expectation of life due to preventable deaths may well be much larger than his largest value of 16.2 years.

It should be emphasized that our estimate of 12.67 years as the loss in the average length of life, due to preventable deaths, does not refer to the average person irrespective of age, but only to those at the age of birth. Professor Fisher's estimate of from 14 to 16 years upon the whole of life is the only estimate he can make by his short-cut method, and it is very easy for the average person to fall into the error indicated. The loss at any age is significant enough, but by far the greatest loss is sustained at the age of birth.

On reference to Table 6, we notice that even at age 10 the loss in expectation of life is 8.15 years, or very little over 60 per cent. of the loss at the age of birth. At age 40, the loss is a little over 5 years, or about 40 per cent. of the loss at the age of birth. The loss at each age higher than age 10 is approximately 1 year less than the loss for age 10, for each succeeding ten years.

For purposes of comparison, attention is called to the fact that it has been estimated elsewhere by means of the methods used in this paper,* that the occurrence of deaths due to

* J. W. Glover, "The Monetary Loss in the United States due to Tuberculosis, based on the Returns of the Twelfth Census," *Transactions* of the Sixth International Congress on Tuberculosis, 1908.

tuberculosis causes a loss of about 2½ years in the expectation of life at age 20, or about one third the loss due to preventable deaths at the same age. However, this estimate is based upon the total number of deaths due to tuberculosis instead of just the preventable deaths due to that disease.

Similarly, the loss in the average length of life due to the total number of deaths due to typhoid fever has been estimated elsewhere to be about one half of one year, or about one twenty-fifth of the loss due to preventable deaths at the same age (age of birth).*

The important question naturally arises, would an assumed prevention of deaths less conservative in character than the one we have just considered lead to a little or to a much greater variation in estimates of vital losses due to preventable deaths? In other words, once we have actually decreased the number of annual deaths due to the different diseases to accord with Professor Fisher's ratios, have we accomplished all that is worth while, or have we merely begun? We shall attempt to answer these questions by the use of deductions made in connection with a discussion of Table 3.

It is to be remembered that Table 3 was constructed with death rates based upon deaths due only to unpreventable diseases. Thus, Table 3 reflects the mortality conditions of a community wherein occur no deaths due to diseases of which any percentage of deaths are at present preventable.

It is true that the assumption of such a community is too ideal to be realized for a long time in the future, but we believe this assumption can be replaced by others, less extreme in character, which would lead to results almost as remarkable. If Professor Fisher had tabulated the most radical or extreme estimates of the ratios of preventability that were given him, instead of the most conservative average of all, and if absolutely everything in the form of recent knowledge and discoveries in regard to diseases were used, we believe the results given in Table 3 would at least be approached very closely.

By the nature of the facts involved, it would be impossible to completely verify the above statement; the belief is based

* W. C. Mendenhall and Earl W. Castle, "Vital and Monetary Losses in the United States due to Typhoid Fever, QUARTERLY PUBLICATION OF THE AMERICAN STATISTICAL ASSOCIATION, June, 1911.

solely upon a personal survey of all the data used in the preparation of this paper and we shall not take the time or space to attempt to substantiate it.

In order that the main features of Table 3 might be more clearly and readily comprehended, we have brought together in Table 7 representative values of the death rates and expectations of life of both Tables 1 and 3. Columns of corresponding differences and percentages are added.

TABLE 7.

Age.	(Table 1).	(Table 3).	Differences.	Percentages.	(Table 1).	(Table 3).	Differences.	Percentages.
(1)	(2)	(3)	(4)	(4)÷(2)	(2)′	(3)′	(4)′	(4)÷(2)′
0.....	.134300	.006908	.127392	95.1	49.44	84.89	35.45	71.8
10.....	.002407	.000058	.002349	97.5	51.57	75.55	23.98	46.5
20.....	.005222	.000084	.005138	98.4	43.15	65.60	22.45	52.0
30.....	.007554	.000242	.007312	96.8	35.61	55.68	20.07	56.4
40.....	.009631	.000732	.008899	92.4	28.30	45.88	17.58	62.1
50.....	.014501	.001725	.012776	88.2	21.17	36.36	15.19	71.7
60.....	.028729	.003630	.025099	87.5	14.63	27.14	12.51	85.7
70.....	.057337	.007646	.049691	86.7	9.29	18.25	8.96	96.4
80.....	.12638	.02321	.10317	82.0	5.32	9.96	4.64	87.2

It is not our purpose to discuss all the results indicated by Table 7, but merely to dwell briefly upon the most important results which will help us most to answer the questions raised at the beginning of this discussion.

According to Table 7, the average length of life should be about 85 years, or about 35½ years longer than it is at the present time; expressed differently, the average length of life should be about 70 per cent. longer than it is at present.

The expectation of life at age 10 should be about 76 years, or 24 years longer than it is at present.

We can now answer our very important questions by saying that the results given in Table 7 indicate that the results obtained by comparing Tables 1 and 2 are little more than marks of a beginning of what can be done in the way of improving general health conditions.

Such results at least suggest excellent goals toward which every effort should be made to advance; the results are certainly worthy of the greatest efforts.

MONETARY LOSSES.

Our discussion of monetary losses is based primarily upon the assumption that the average person contributes something to the wealth of the community about him during the most productive period of his life.

Although practically no one will object to the above assumption, there seems to be no satisfactory way of estimating the average value of this contribution. There are many who not only contribute nothing, but are even a heavy expense to their community all during their lives; there are others whose contribution is immeasurably great. There are some who, because of ill health or poverty, are prevented from adding to the wealth of their community except at particular intervals of the period throughout which the average person is considered most productive. Only a systematic investigation into these various sets of conditions would give us any satisfactory results, and such an investigation seems never to have been made.

Many have constructed schedules of values of this contribution by ages, which are intended to fit the average set of circumstances; the fact that these schedules fail to agree very closely indicates that they are not altogether satisfactory.

We make no attempt in this paper to decide rigidly what the value of this contribution should be assumed to be; instead, we assume a purely arbitrary value and then discuss certain monetary losses in terms of this value. Those who have decided views as to what the value of the contribution should be assumed to be can then modify the results in exactly the same proportion that their choice of this value differs from the value used here.

We shall assume that the average person contributes $100 annually to the wealth of the community about him between the ages 20 and 70. Hence, the death of such a person involves a loss equal to the present value of an annuity of $100 per annum at his age for the period ending with the age 70.

The present values of such annuities for each age from 20 to 70, interest at 5 per cent., based upon Tables 1 and 2, are given in the following table:

TABLE 9.

PRESENT VALUE, AT EACH AGE FROM 20 TO 70, COMPUTED WITH COMPOUND INTEREST AT 5 PER CENT., OF AN ANNUITY OR WEALTH INCREMENT OF $100 PER ANNUM AT THE END OF EACH YEAR, UNTIL AGE 70, ACCORDING TO TABLES 1 AND 2.

Age.	Deaths		Loss in Value.	Age.	Deaths		Loss in Value.
	Not Prevented.	Prevented.			Not Prevented.	Prevented.	
20	$1,601.00	$1,705.66	$104.66	45	$1,189.61	$1,274.46	$84.85
21	1,589.88	1,695.02	105.14	46	1,163.94	1,247.11	83.17
22	1,578.45	1,683.97	105.52	47	1,137.51	1,218.79	81.28
23	1,566.72	1,672.51	105.79	48	1,110.16	1,189.40	79.24
24	1,554.75	1,660.64	105.89	49	1,081.68	1,158.83	77.15
25	1,542.58	1,648.38	105.80	50	1,051.89	1,126.97	75.08
26	1,530.17	1,635.69	105.52	51	1,020.73	1,093.73	73.00
27	1,517.49	1,622.55	105.06	52	988.11	1,059.05	70.94
28	1,504.49	1,608.95	104.46	53	954.13	1,022.92	68.79
29	1,491.10	1,594.83	103.73	54	918.94	985.43	66.49
30	1,477.25	1,580.19	102.94	55	882.76	946.67	63.91
31	1,462.93	1,564.99	102.06	56	845.57	906.58	61.01
32	1,448.11	1,549.22	101.11	57	807.37	865.14	57.77
33	1,432.74	1,532.85	100.11	58	768.09	822.25	54.16
34	1,416.76	1,515.81	99.05	59	727.38	777.64	50.26
35	1,400.11	1,498.04	97.93	60	684.89	731.03	46.14
36	1,382.71	1,479.50	96.79	61	640.40	682.21	41.81
37	1,364.55	1,460.16	95.61	62	593.66	630.93	37.27
38	1,345.58	1,440.00	94.42	63	544.33	576.92	32.59
39	1,325.82	1,418.99	93.17	64	492.12	519.91	27.79
40	1,305.25	1,397.16	91.91	65	436.69	459.16	22.92
41	1,283.84	1,374.40	90.56	66	377.52	395.60	18.08
42	1,261.55	1,350.77	89.22	67	314.02	327.44	13.39
43	1,238.39	1,326.27	87.88	68	245.51	254.50	8.99
44	1,214.41	1,300.84	86.43	69	171.13	176.18	5.05

With these present values as a basis we could compute the present value of the total losses sustained throughout the United States, if the number of deaths by causes and by ages were known for the whole country. Less than one half of our states, however, keep anything like an accurate record of deaths each year.

If these statistics of deaths were known for the whole country, we could use Professor Fisher's table of ratios to determine the number of deaths at each age that are preventable; this number multiplied by the present value of the corresponding annuity would give us a measure of the monetary loss for that age, due to the occurrence of preventable deaths.

Since these statistics are not available for the whole country, we are compelled to use another method to determine the monetary losses due to deaths that are preventable. We shall make use of the population instead of the number of deaths to determine the losses under consideration, because the population of the whole country is given every 10 years by the census.

Just as the second column in Table 9 gives the present values of the future contributions, as computed by the use of Table 1, the third column gives the present values of the same future contributions, as computed by the use of Table 2 which is based upon mortality conditions which would exist if preventable deaths were prevented.

Referring to Table 9, we see that the value of the average person aged 20 to his community is $1,601.00, and that this value would be $1,705.66 if preventable deaths were prevented. Hence, such a community suffers a loss of $104.66 on every person living at the age 20, because of the occurrence of preventable deaths.

Column four of Table 9 gives values of these differences between the present values of persons for the different ages under the two sets of mortality conditions, which may therefore be regarded as measures of the effect of preventable deaths upon the value of persons to their community at different ages.

These differences, or "loss rates" increase slightly at first as age increases, due to the peculiar combination of the expectation of life and number of survivors of the mortality tables at these ages. After an interval of about ten years the loss rates start to diminish very gradually until about age 50 and then the decrease becomes more rapid toward zero.

Since we have a monetary measure of the effects of the occurrence of preventable deaths upon the value of each living person to his community, it remains merely to multiply the population of the United States for each age by the corresponding loss rate to determine the total losses at each age due to the occurrence of preventable deaths.

These populations for 1910 are given in Table 10 together with the corresponding loss rates, and finally the total losses

themselves. The population data were obtained in quinquennial age groups from the government mortality statistics, and the population for each age was determined by Sprague's method, discussed and used previously in this paper.

TABLE 10.

TABLE SHOWING THE PRESENT VALUE, COMPOUNDED ANNUALLY AT 5 PER CENT. OF THE LOSS DUE TO PREVENTABLE DEATHS, BASED UPON THE POPULATION OF THE UNITED STATES OF 1910, BOTH MALES AND FEMALES, FOR EACH AGE AND CERTAIN AGE GROUPS BETWEEN AGES 20 AND 70, ON THE BASIS OF AN ASSUMED PRODUCING CAPACITY OF $100 PER ANNUM UNTIL AGE 70.

Age.	Population, 1910.	Loss Rate.	Total Loss.	Age.	Population, 1910.	Loss Rate.	Total Loss.
20.....	1,829,028	104.66	$191,426,070	45.....	947,320	84.85	$80,380,102
21.....	1,834,552	105.14	192,884,797	46.....	915,046	83.17	76,104,376
22.....	1,827,619	105.52	192,850,357	47.....	887,879	81.28	72,166,805
23.....	1,802,149	105.79	190,649,343	48.....	867,715	79.24	68,757,737
24.....	1,763,636	105.89	186,751,416	49.....	851,237	77.15	65,672,935
20-24.....	9,056,984		954,561,983	20-49.....	40,336,056		3,960,550,043
25.....	1,725,192	105.80	182,525,314	50.....	835,059	75.08	62,696,230
26.....	1,686,044	105.52	177,911,363	51.....	822,047	73.00	60,009,431
27.....	1,641,354	105.06	173,340,651	52.....	796,738	70.94	56,520,594
28.....	1,590,635	104.46	166,157,732	53.....	751,796	68.79	51,716,047
29.....	1,536,778	103.73	159,409,982	54.....	695,151	66.49	46,220,590
20-29.....	17,236,987		1,813,907,025	20-54.....	44,236,847		4,237,612,935
30.....	1,481,263	102.94	152,481,213	55.....	641,014	63.91	40,967,205
31.....	1,421,512	102.06	145,079,515	56.....	585,274	61.01	35,707,567
32.....	1,375,605	101.11	139,087,422	57.....	541,109	57.77	31,259,867
33.....	1,351,976	100.11	135,346,317	58.....	516,108	54.16	27,952,409
34.....	1,341,829	99.05	132,908,162	59.....	503,446	50.26	25,303,196
20-34.....	24,209,172		2,517,809,654	20-59.....	47,023,798		4,397,803,179
35.....	1,329,384	97.93	130,186,575	60.....	488,856	46.14	22,555,816
36.....	1,319,524	96.79	127,716,728	61.....	475,366	41.81	19,875,052
37.....	1,297,068	95.61	124,012,671	62.....	458,783	37.27	17,098,842
38.....	1,253,418	94.42	118,347,728	63.....	435,645	32.59	14,197,671
39.....	1,196,706	93.17	111,497,098	64.....	408,500	27.79	11,352,215
20-39.....	30,605,272		3,128,570,454	20-64.....	49,290,948		4,482,882,775
40.....	1,143,469	91.91	105,096,236	65.....	383,517	22.92	8,790,210
41.....	1,090,171	90.56	98,725,886	66.....	359,649	18.08	6,502,454
42.....	1,043,061	89.22	93,061,902	67.....	335,775	13.39	4,496,027
43.....	1,006,836	87.88	88,480,748	68.....	312,015	8.99	2,805,015
44.....	978,050	86.43	84,532,862	69.....	288,547	5.05	1,457,162
20-44.....	35,866,859		3,597,468,088	20-69.....	50,970,451		4,505,933,543

As in the case of the loss rates, the total monetary losses for the different ages increase slightly at first as age increases from about $191,000,000 at age 20 to $193,000,000 at age 22, but begin to decrease a little earlier than the loss rates. At age 23 the loss is again approximately that at age 20; from age 23 on, the decrease in the values of the losses is very gradual to the end of the table.

Besides the total losses for each age, the losses for the accumulative age groups 20–24, 20–29,—20–69 are given. The grand total, or the loss for the total ages 20–69 is $4,505,933,543. This amount may then be regarded as the total monetary loss in terms of the arbitrarily assumed contribution $100, sustained by this country because of the occurrence of unnecessary deaths.

Just as in Table 9, we have the present value, at each age, of the future contributions of $100 per annum, so we have in this grand total of $4,505,933,543 the present value of an annuity of $246,820,350 per annum to continue for 50 years, interest at 5 per cent. In other words, $246,820,350 may, in the same connection, be regarded as the total annual loss sustained in this country because of preventable deaths. Perhaps it will help to visualize the value of this annual loss to state that it is about one fourth the value of our annual wheat crop and between two and three times the value of our annual product of gold which is approximately one fifth of that of the world.

A comparison of the mortality tables in this article leads to the following conclusions.

The losses in death rates sustained by this country due to preventable deaths run from as low as 1 per 1,000 at age 10 to as high as 57 per 1,000 at the age of birth. Expressed in percentages, these losses run from 30 per cent. of the actual death rates, or less, at age 90 and above, to over 50 per cent. at age 35 and all ages below, being over 55 per cent. between ages 15 and 20.

The losses in expectation of life or average future life time are about 13 years on the whole of life, 5 years less than that at the age of 10, and approximately 1 year less than that for each succeeding ten years.

If we assume that every person contributes $100 annually to the wealth of the community about him during the most productive period of life, say from age 20 to age 70, we estimate the present value of the total future losses in this country to be about $4,500,000,000 and the annual loss to be about $250,000,000.

The estimates of corresponding monetary losses where a different value is assumed for the individual contribution are to be found by increasing or diminishing (as the case may be) the above values in the same proportion.

JOINT COMMITTEE ON STANDARDS FOR GRAPHIC PRESENTATION.

PRELIMINARY REPORT PUBLISHED FOR THE PURPOSE OF INVITING SUGGESTIONS FOR THE BENEFIT OF THE COMMITTEE.*

As a result of invitations extended by The American Society of Mechanical Engineers, a number of associations of national scope have appointed representatives on a Joint Committee on Standards for Graphic Presentation. Below are the names of the members of the committee and of the associations which have coöperated in its formation.

Willard C. Brinton, *Chairman*, American Society of Mechanical Engineers. 7 East 42d Street, New York City.
Leonard P. Ayres, *Secretary*, American Statistical Association. 130 East 22d Street, New York City.
N. A. Carle, American Institute of Electrical Engineers.
Robert E. Chaddock, American Association for the Advancement of Science.
Frederick A. Cleveland, American Academy of Political and Social Science.
H. E. Crampton, American Genetic Association.
Walter S. Gifford, American Economic Association.
J. Arthur Harris, American Society of Naturalists.
H. E. Hawkes, American Mathematical Society.
Joseph A. Hill, United States Census Bureau.
Henry D. Hubbard, United States Bureau of Standards.
Robert H. Montgomery, American Association of Public Accountants:
Henry H. Norris, Society for the Promotion of Engineering Education.
Alexander Smith, American Chemical Society.
Judd Stewart, American Institute of Mining Engineers.
Wendell M. Strong, Actuarial Society of America.
Edward L. Thorndike, American Psychological Association.

*Copies may be had from the American Society of Mechanical Engineers, 29 West 39th St., New York. Price, 10 cents. Discount on quantities.

The committee is making a study of the methods used in different fields of endeavor for presenting statistical and quantitative data in graphic form. * * * * If simple and convenient standards can be found and made generally known, there will be possible a more universal use of graphic methods with a consequent gain to mankind because of the greater speed and accuracy with which complex information may be imparted and interpreted.

THE FOLLOWING ARE SUGGESTIONS WHICH THE COMMITTEE HAS THUS FAR CONSIDERED AS REPRESENTING THE MORE GENERALLY APPLICABLE PRINCIPLES OF ELEMENTARY GRAPHIC PRESENTATION

1 The general arrangement of a diagram should proceed from left to right.

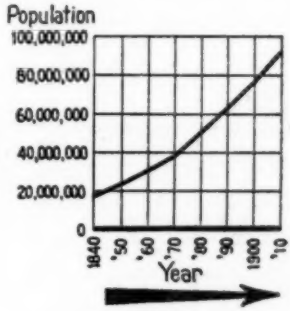


Fig. 1

Fig. 2

2 Where possible represent quantities by linear magnitudes as areas or volumes are more likely to be misinterpreted.

3 For a curve the vertical scale, whenever practicable, should be so selected that the zero line will appear on the diagram.

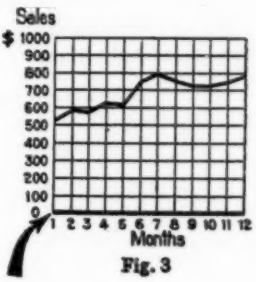


Fig. 3

4 If the zero line of the vertical scale will not normally appear on the curve diagram, the zero line should be shown by the use of a horizontal break in the diagram.

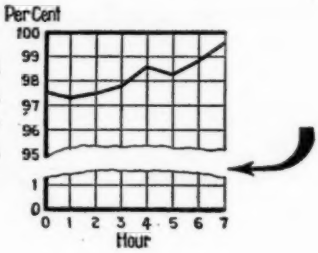


Fig. 4

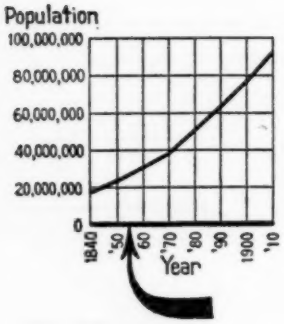


Fig. 5A

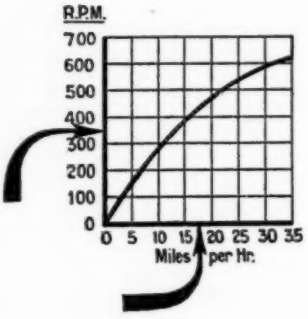


Fig. 5B

5 The zero lines of the scales for a curve should be sharply distinguished from the other coördinate lines.

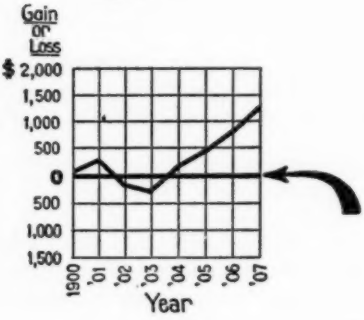


Fig. 5C

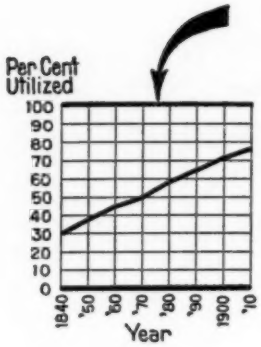


Fig. 6A

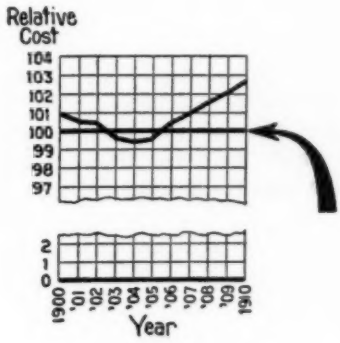


Fig. 6B

6 For curves having a scale representing percentages, it is usually desirable to emphasize in some distinctive way the 100 per cent line or other line used as a basis of comparison.

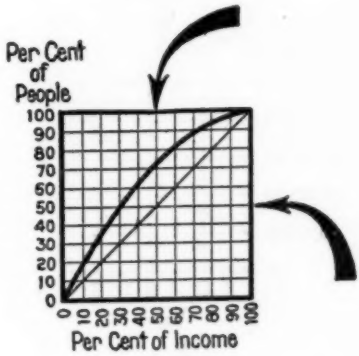


Fig. 6C

7 When the scale of a diagram refers to dates, and the period represented is not a complete unit, it is be ter not to emphasize the first and last ordinates, since such a diagram does not represent the beginning or end of time.

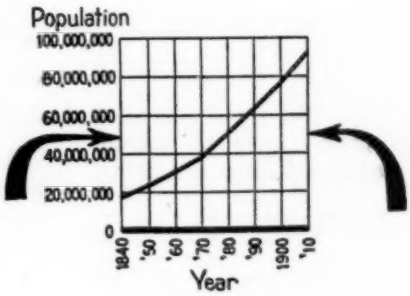


Fig. 7

8 When curves are drawn on logarithmic coördinates, the limiting lines of the diagram should each be at some power of ten on the logarithmic scales.

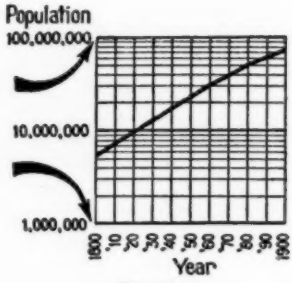


Fig. 8

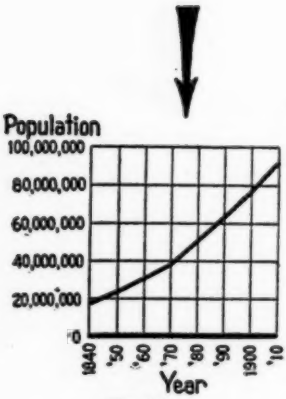


Fig. 9A

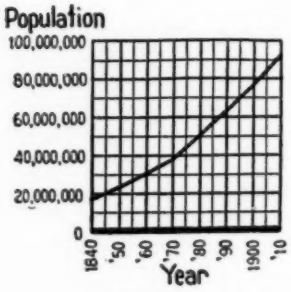


Fig. 9B

9 It is advisable not to show any more coördinate lines than necessary to guide the eye in reading the diagram.

10 The curve lines of a diagram should be sharply distinguished from the ruling.

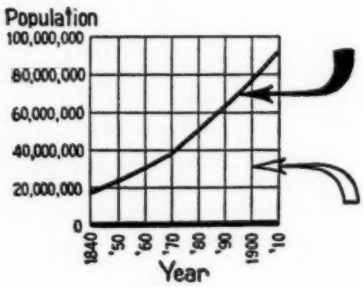


Fig. 10

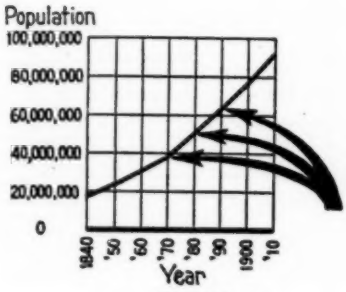


Fig. 11A

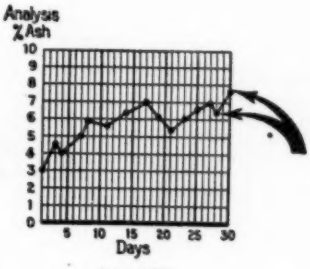


Fig. 11B

11 In curves representing a series of observations, it is advisable, whenever possible, to indicate clearly on the diagram all the points representing the separate observations.

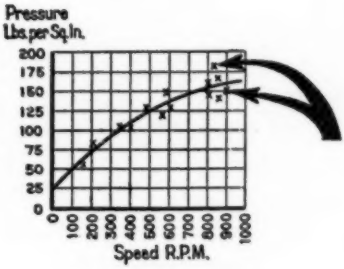


Fig. 11C

12 The horizontal scale for curves should usually read from left to right and the vertical scale from bottom to top.

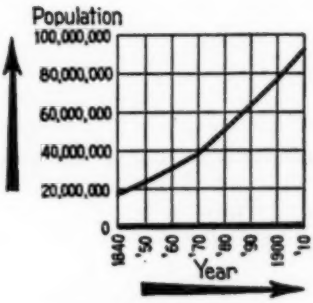


Fig. 12

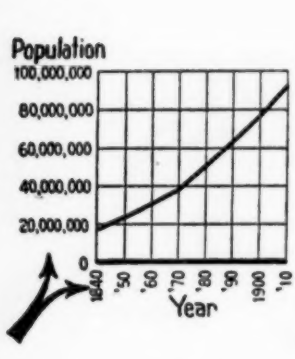


Fig. 13A

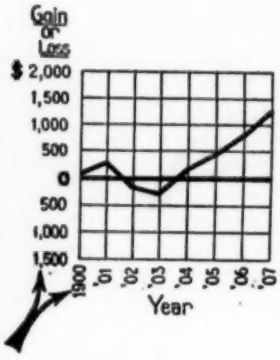


Fig. 13B

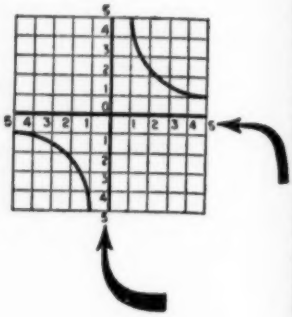


Fig. 13C

13 Figures for the scales of a diagram should be placed at the left and at the bottom or along the respective axes.

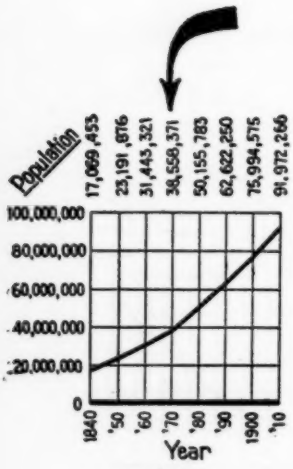


Fig. 14A

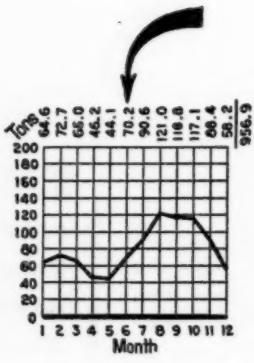


Fig. 14B

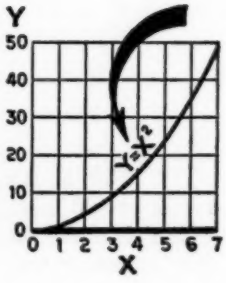
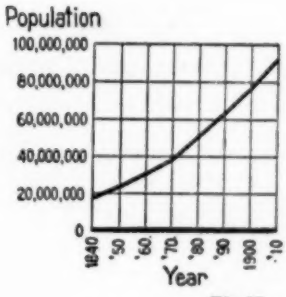


Fig. 14C

14 It is often desirable to include in the diagram the numerical data or formulae represented.

15 If numerical data are not included in the diagram it is desirable to give the data in tabular form accompanying the diagram.



Year	Population
1840	17,069,453
1850	23,191,876
1860	31,443,321
1870	38,558,371
1880	50,155,783
1890	62,622,250
1900	75,994,575
1910	91,972,266

Fig. 15

16 All lettering and all figures on a diagram should be placed so as to be easily read from the base as the bottom, or from the right-hand edge of the diagram as the bottom.

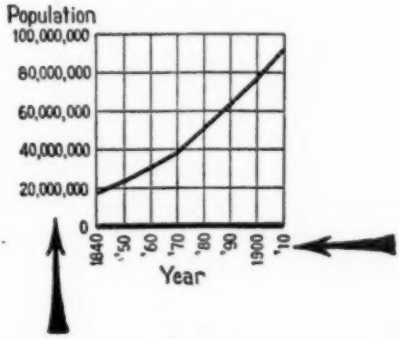
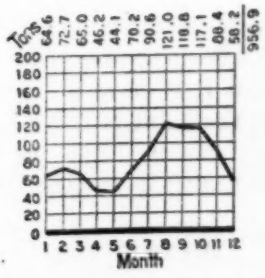


Fig 16

17 The title of a diagram should be made as clear and complete as possible. Sub-titles or descriptions should be added if necessary to insure clearness.



Aluminum Castings Output of Plant No. 2, by Months, 1914.

Output is given in short tons.

Sales of Scrap Aluminum are not included.

Fig. 17

NEW METHOD FOR COMPUTING THE MOVING AVERAGE.

By Willford I. King, Ph.D., *University of Wisconsin.*

Ability to readily use modern statistical methods has come to be an almost essential prerequisite to original research in the fields of economics and sociology. These sciences both deal largely with historical data. In studying the oscillations of historical variables it is nearly always necessary to compare them with some kind of a trend, and no other method of locating the trend is so generally applicable as the use of the moving average.

The usual methods of computing the moving average are satisfactory enough when the number of items in each group is small, but when a large group is used the labor involved becomes so great and the opportunities for error are so numerous that statisticians have been discouraged from employing this most useful tool. By the following method, the moving average can be computed with little extra work, even though the number of items in the group is largely increased. Each operation is checked for the detection of errors. An adding-machine is evidently a prime necessity if moving averages are to be calculated in this manner.

The method is believed to be original, though it is so simple that it is entirely possible that many other statisticians are in the habit of using it. The writer has found it convenient and practical and the students in his classes have had little difficulty in putting the device into practice.

Method for Computing the Moving Average.

1. Determine upon the correct number of items constituting one wave-length.

2. By aid of the adding-machine, summate the items constituting the first wave-length; *e. g.*, if there are seven items in one wave-length, add the first seven items, sub-totalling at the close. Continue adding, one by one, the succeeding items

of the variable, sub-totalling after *each* item is inserted. This is shown in Column I of the illustrative table.

3. Next, begin by inserting the first item of the variable in the adding-machine. Sub-total. Add the succeeding items of the variable, sub-totalling after *each* item. The spacing on the adding-machine slip must be identical with that obtained in the operation recorded in paragraph 2. The final result is illustrated in Column II.

4. Shift the second adding-machine slip downward until the first item thereon falls opposite the first item of the *second* wave-length in the first column. This is shown in Column II of the illustrative table where the *first* item is placed opposite the *eighth* item in Column I. With the adding-machine slips in this position, paste the second to the first.

5. At the head of a third column and directly to the right of its original position, enter the first sub-total of Column I. Now subtract each sub-total in the second column from the adjacent sub-total in the first column and enter the remainders immediately to the right in the third column. That part of Column II which extends below Column I is discarded.

6. Add each sub-total in Column II to the adjacent quantity in Column III. The sums, if correct, check with the sub-totals in Column I.

7. Divide each of the remainders in the third column by the number of items composing one group or wave. The quotients are the items of the moving-average. The date for the first of these items corresponds to the date of the middle item of the first wave-length in the original variable. For example, in Column I of the illustrative table, the date of the fourth or middle item of the first wave is 1893, which, therefore, is likewise the date of the first item in Column IV.

8. Check the mathematical accuracy of the last operation by summating Column IV, multiplying the sum by the number of items in one wave-length, and comparing with the sum of Column III. In case of discrepancy, the columns may be divided by horizontal lines into segments and each segment may be separately checked in the above manner until the location of the error is discovered.

AN ILLUSTRATION OF THE COMPUTATION OF A SEVEN-YEAR MOVING AVERAGE BY THE NEW PROCESS

Column I		Column II		Column III		Column IV
Pig Iron Production in the United States.		Pig Iron Production in the United States.		Column I—Column II		Moving Average.
Year.	Thousands of Tons.	Year.	Thousands of Tons.	Year.	Remainders.	Remainders ÷ Seven.
1890	*9,203*					
1891	*8,280*					
1892	*9,157*					
1893	*7,125*					
1894	*6,658*					
1895	*9,446*					
1896	*8,623*					
	58,492			1893	**58,492**	**8,356**
1897	*9,653*	1890	*9,203*			
	68,145		**9,203**	1894	**58,942**	**8,420**
1898	*11,774*	1891	*8,280*			
	79,919		**17,483**	1895	**62,436**	**8,919**
1899	*13,621*	1892	*9,157*			
	93,540		**26,640**	1896	**66,900**	**9,557**
1900	*13,789*	1893	*7,125*			
	107,329		**33,765**	1897	**73,564**	**10,509**
1901	*15,878*	1894	*6,658*			
	123,207		**40,423**	1898	**82,784**	**11,826**
1902	*17,821*	1895	*9,446*			
	141,028		**49,869**	1899	**91,159**	**13,023**
1903	*18,009*	1896	*8,623*			
	159,037		**58,492**	1900	**100,545**	**14,364**
1904	*16,497*	1897	*9,653*			
	175,534		**68,145**	1901	**107,389**	**15,341**
1905	*22,992*	1898	*11,774*			
	198,526		**79,919**	1902	**118,607**	**16,944**
1906	*25,307*	1899	*13,621*			
	223,833		**93,540**	1903	**130,293**	**18,613**
1907	*25,781*	1900	*13,789*			
	249,614		**107,329**	1904	**142,285**	**20,326**
1908	*15,936*	1901	*15,878*			
	265,550		**123,207**	1905	**142,343**	**20,335**
1909	*25,795*	1902	*17,821*			
	291,345		**141,028**	1906	**150,317**	**21,474**
1910	*27,304*	1903	*18,009*			
	318,649		**159,037**	1907	**159,612**	**22,802**
1911	*23,650*	1904	*16,497*			
	342,299		**175,534**	1908	**166,765**	**23,824**
1912	*29,727*	1905	*22,992*			
	372,026		**198,526**	1909	**173,500**	**24,786**
1913	*30,966*	1906	*25,307*			
	402,992		**223,833**	1910	**179,159**	**25,594**
1914	*23,332*	1907	*25,781*			
	426,324		**249,614**	1911	**176,710**	**25,244**
				Total ..	**2,241,802**	**320,257**

COMPARATIVE MILITARISM.

By Arthur MacDonald, *Anthropologist, Washington, D. C.*

One way to estimate militarism in a country is to give the number of soldiers and sailors (including officers) in time of peace, relative to the population, which is done in column 2 of the following table. The countries are arranged according to a descending scale of militarism, Roumania being highest, which is 15,700 soldiers and sailors per million population. United States is last having only 1,111 soldiers and sailors per million population. Sweden is second, having 15,299 soldiers and sailors per million population. France is third, having 14,874 per million population, but she is highest of all the larger nations in degree of militarism. Servia is fourth, Bulgaria fifth, and Turkey sixth. Germany is seventh, but of the larger nations she is second, being next to France in degree of militarism. Italy is eighth, Greece ninth, Austria-Hungary tenth, and so on down.

In some of the countries no trustworthy data as to sailors are obtainable, especially in the Balkan States, which, however, stand very high in militarism on land; they are rated without regard to their militarism on sea. If we compare the larger nations only, the order in degree of militarism (column 2) is France first, then Germany, Italy, Austria-Hungary, Russia, Great Britain, and the United States which is by far the least military of all.

In column 3, the figures are given for militarism on land only; the order is practically the same as in column 2 for both land and sea, except Great Britain, which is very low in militarism on land. In column 4, the figures are given for militarism on sea alone, the order being Great Britain first, then Sweden, Italy, France, Germany, Russia and so on. Great Britain's militarism on sea is more than three times as much as any other country; she is the only country having more sailors than soldiers. It will be found on the whole that militarism is according to geographical and political necessity.

In columns 5 and 6 are given the absolute numbers of soldiers and sailors (officers and men) in army and navy. The popular mind generally rates the militarism of a country according to the size of its army, which is incorrect, for the army might be large simply because the country is large; thus Russia has by far the largest army; so the United States, a non-military country, has a larger army than most of the Balkan States, which are bristling with militarism.

Also it would not do to estimate militarism according to amount of money paid out per capita of population (columns 7 and 8); for a country might have a much higher degree of militarism than it had money to carry it out; also larger and richer nations would have a high degree of militarism. Thus Great Britain's per capita rate for army and navy combined is the highest, making it most military of all, which is not true.

In column 9 is given the number of bordering nations for each country. Other things being equal, a nation having a large number of countries bordering upon it would be expected to have a larger army for its protection, especially if these other countries were first-class powers.

The true way, therefore, to estimate militarism, is according to the number of both soldiers and sailors, relative to population in times of peace.

COMPARATIVE MILITARISM.*

(Time of Peace.)

Countries 1911 (Population)	Number of Officers and Men per Million Population in both Army and Navy.	Number of Officers and Men per Million Population.		Absolute Number of Officers and Men.		Cost per Capita of Population.		Number of Bordering Countries
		Army.	Navy.	Army.	Navy.	Army.	Navy.	
1	2	3	4	5	6	7	8	9
Roumania (5,956,690)	15,700	15,700	—	98,139	—	$2.15	$—	4
Sweden (5,521,943)	15,299	14,400	899	79,603	4,965	2.43	0.95	2
France (39,601,509)	14,874	14,231	643	563,596	25,500	3.90	2.00	5
Servia (2,911,701)	12,220	12,220	—	35,605	—	1.47	—	6
Bulgaria (4,329,108)	11,041	11,041	—	57,800	—	1.83	—	3
Turkey (35,400,000)	10,593	10,593	—	375,000	—	1.00	—	5
Germany (64,903,423)	9,683	9,167	516	626,732	33,500	3.00	1.70	7
Italy (34,686,683)	9,273	8,397	876	291,293	30,398	2.28	1.12	3
Greece (2,263,952)	8,966	8,966	—	23,600	—	1.40	—	1
Austria-Hungary (49,161,766)	8,359	8,057	302	396,114	14,888	1.90	0.20	7
Russia (118,690,600)	7,660	7,300	360	1,200,000	60,000	1.50	0.31	6
Great Britain (45,370,530)	6,963	2,953	4,010	134,000	181,939	2.50	4.50	0
Belgium (7,516,730)	6,100	6,100	—	46,574	—	1.74	—	3
Netherlands (5,945,155)	6,012	5,830	182	34,662	10,822	2.00	—	2
United States (91,972,266)	1,111	982	91	91,783	12,032	1.00	1.35	2

* The figures in this table are based upon official data. Blanks indicate that reliable data were not obtainable.

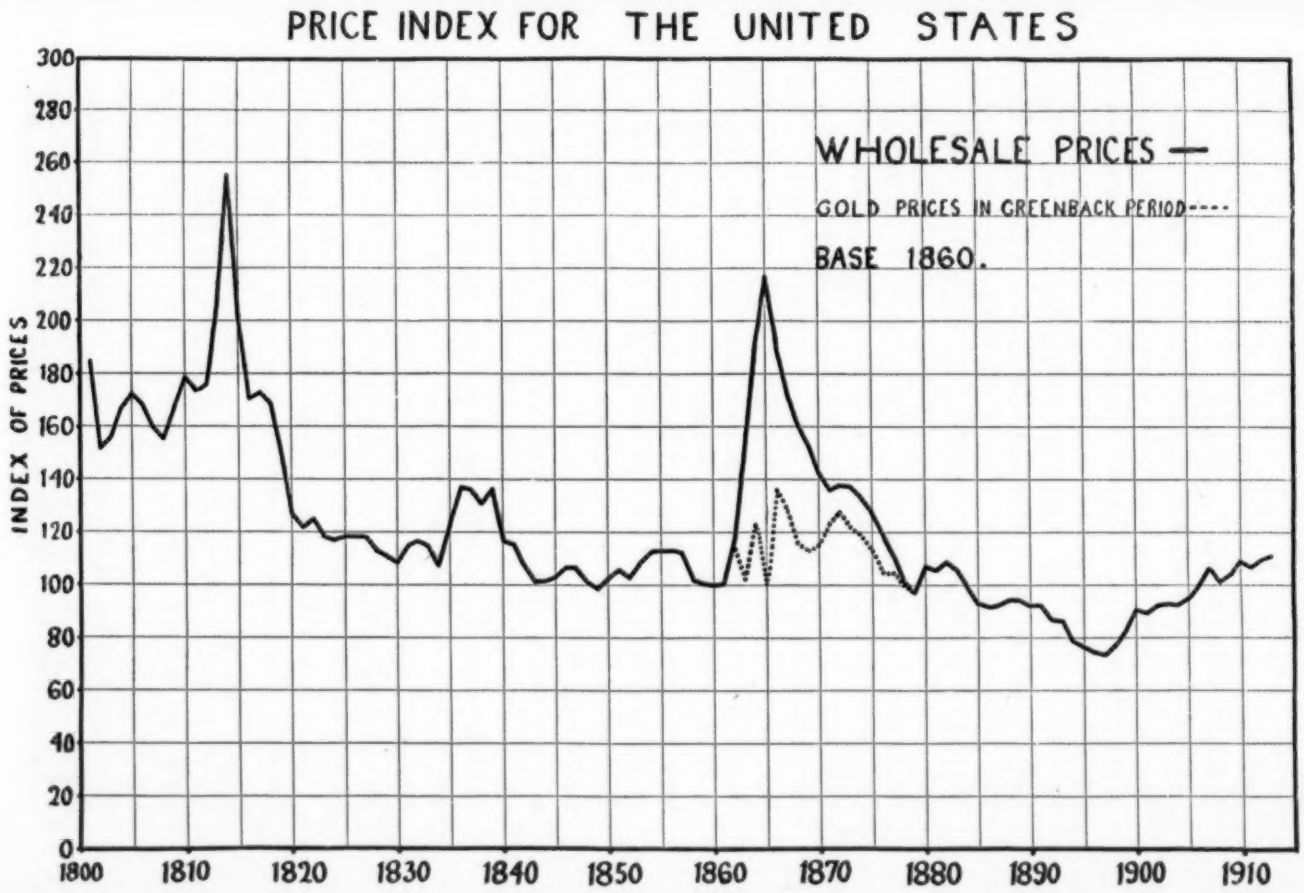
WHOLESALE PRICES FOR THE UNITED STATES, 1801–1840.

By Alvin H. Hansen.

There has not been, so far as the writer is aware, any study of prices for the United States prior to 1825, and very little prior to 1840. Since the trend of prices is a subject of great interest and importance in the study of American social and economic movements, this investigation has been undertaken in order to construct a curve of prices for the United States extending back to the beginning of the last century.

The price curve shown in the accompanying chart is a composite of three separate index series all having been reduced to the common base 1860. The first period, 1801–1840, was worked up by the writer, the price average for the year 1825 being taken as the original base. The figures for the period 1840–1891 are the simple average index numbers prepared for the Aldrich Report by Roland P. Falkner, 1860 being the base. For the period 1891–1913 the Bureau of Labor index numbers were used, based 1890–9.

The numbers for the period 1801–1840 were derived from two distinct series, both having the year 1825 as the base. The index numbers for 1801–1825 were derived from the monthly quotations of wholesale prices of 79 commodities quoted in the Boston *Gazette* 1801–1815, and in the Boston *Patriot* 1816–1825. The index numbers for 1825–1840 are based on the monthly quotations of New York wholesale prices of 63 commodities quoted in the Report of the Secretary of the Treasury for 1863. Where various grades of unimportant articles were quoted, only one was used. When a range of a higher and lower price was shown in the quotations, the mean or average price was found. The average annual price for each commodity was obtained by adding the monthly quotations, taken for the first of the month or as close thereto as possible, and dividing the sum by the number of months. In some cases it was not possible to get quotations for each month. No quotations at all could be secured for 1809. For


PRICE INDEX FOR THE UNITED STATES
WHOLESALE PRICES —
GOLD PRICES IN GREENBACK PERIOD----
BASE 1860.
INDEX OF PRICES
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910

1825–1840 the yearly averages quoted in the Report of the Secretary of the Treasury were used.

Index numbers were worked out for each separate commodity, the price for 1825 being used as the base for both periods under consideration. That is to say, the average annual price for each year was reduced to a percentage of the average price for 1825. In the case of each commodity the same grade was used throughout, with the exception of iron. In this case "pig iron" was used for 1801–1808, "Philadelphia bar" for 1808–1815, and "Russia" for 1815–1825. The two latter sold at very nearly the same price, so "Russia" was simply substituted for "Philadelphia bar." The prices of these two and that of pig iron were, however, so different that they had to be adjusted. It was assumed that the index numbers of pig iron would indicate in general the trend of prices of "Philadelphia bar" and "Russia." The annual prices of pig iron were therefore reduced to index numbers on the base 1808, and the prices of "Philadelphia bar" and "Russia" were reduced to index numbers on the base 1825. The index number of the latter variety in 1808 was 111.9, while that of pig iron would be 100.0. Each of the index numbers of pig iron back to 1801 was, therefore, multiplied by 111.9 in order to show the general trend of iron prices in the entire period. In the case of the other 78 commodities used for the period 1801–1825, it was possible to carry through the same grade throughout. The attempt was made to carry through 84 commodities, but it was found necessary to throw out 5, because of changes in the quality of the commodity quoted. The same grade was quoted for the 63 commodities used for the period 1825–1840.

The index numbers thus obtained for the different commodities were then summated, and divided by the total number of commodities quoted for that year. The result would be the simple average index number of the prices of all commodities for the given year. Since both periods in question, 1801–1825, and 1825–1840, had the common base, 1825, it is evident that the relative prices of the entire period would thus be established.

The number of months for which quotations were given, and

the number of commodities quoted for each year were as follows:

Source of Information.	Year.	Number of Months for which Quotations were Obtained.	Number of Commodities Quoted.
Boston *Patriot* and Boston *Gazette*.			
	1801	12	75
	1802	12	77
	1803	12	79
	1804	12	79
	1805	12	79
	1805	12	79
	1806	12	79
	1807	7	79
	1808	9	78
	1809	0	0
	1810	12	78
	1811	12	78
	1812	11	78
	1813	11	78
	1814	2	54
	1815	11	78
	1816	7	78
	1817	12	78
	1818	8	79
	1819	11	79
	1820	12	78
	1821	6	78
	1822	12	79
	1823	12	79
	1824	12	79
	1825	12	79
Report of Secretary of Treasury.			
	1825	12	61
	1826	12	62
	1827	12	63
	1828	12	63
	1829	12	63
	1830	12	63
	1831	12	63
	1832	12	63
	1833	12	63
	1834	12	63
	1835	12	63
	1836	12	63
	1837	12	63
	1838	12	62
	1839	12	62
	1840	12	61

The index numbers from 1801 to 1840 are therefore based on the average prices, derived from the monthly quotations, for each year. This method is similar to that employed by the Bureau of Labor in its index, which is also based on the yearly averages of monthly quotations, except in the case of a few of the more unstable commodities, in which cases weekly quotations are used. The Aldrich index numbers, on the other

hand, were based on the actual prices of a given date, January 1, April 1, July 1, or October 1 of a given year. Usually the index numbers were calculated on the basis of January prices in each year, except in the case of commodities for which the January prices would not be the representative prices for the year, as in the case of certain vegetables. In such cases the most representative month was chosen. In a few cases, however, the average prices for the year were used.

The original index numbers, on the base 1825, for the period 1801–1840 were as follows:

1801	155.5	1821	102.3
1802	127.7	1822	105.3
1803	131.2	1823	100.1
1804	140.6	1824	98.9
1805	144.9	1825	100.0
1806	141.6	1826	99.4
1807	134.3	1827	99.6
1808	131.0	1828	95.2
1809	No data	1829	94.1
1810	150.4	1830	91.4
1811	145.7	1831	96.7
1812	148.1	1832	98.1
1813	172.4	1833	96.8
1814	214.8	1834	90.1
1815	168.6	1835	103.8
1816	143.6	1836	115.7
1817	145.4	1837	114.7
1818	141.5	1838	110.2
1819	124.7	1839	115.2
1820	106.9	1840	98.3

In order to get a continuous price curve for the United States for the entire century, the above index was connected up with the Aldrich index. It is recognized that this does not in any way represent an accurate extension of the Aldrich index, since the number and grade of commodities used in the two cases are not identical, and the methods of computation are somewhat different. Nevertheless, such a curve does indicate approximately the general trend of prices.

The method used in connecting up the above index with the Aldrich index was as follows: The relative price, or the index number, for 1840, as shown in the Aldrich Report, was used as the medium through which to compute the index numbers for all the years back to 1801. It was assumed that, had the base been 1860 instead of 1825, the index numbers for the period 1801–1840 would have stood in 1840 at the same relative price level as the index number of the Aldrich Report for the same year. The Aldrich index number for the year 1840

is 116.8. The above index number for 1840 is 98.3. Hence every unit in the above index series was multiplied by $\frac{116.8}{98.3}$ or 1.188. The product would approximately show the relation of the general level of prices of the given year to that of 1860, the year chosen as the base for the Aldrich index.

The following is a table showing the commodities used for the period 1801–1825, 1825–1840, and for 1840, the first year of the Aldrich series. It should be added that the number of commodities used in the Aldrich Report increases from 90 in 1840 to 223 in 1891.

This study of prices in the period 1801–1840 does not purport to be anything more than a tentative beginning. However, as the followin table indicates, the variety and number of commodities used is fairly comparable with those used in the Aldrich Report for the first years. The index numbers are based on the average prices of 12 monthly quotations for each year, in so far as they could be obtained, while the Aldrich numbers are based in general on January prices alone. The Aldrich Report has the advantage of using the price quotations of several cities. Quotations from say Boston, Philadelphia, Charleston, and New Orleans would undoubtedly have been very desirable for our period, particularly because of the greater diversity of wholesale prices then obtaining in different sections of the country, due chiefly to slow and inadequate means of transportation. Such a task would be far too great for one person to undertake.

COMPARATIVE TABLE OF COMMODITIES USED.

Boston *Gazette* and Boston *Patriot*. 1801–1825	Report of Secretary of Treasury. 1825–1840.	Aldrich Report. 1840.
Flour, fine	Flour, wheat, sup.	
	Flour, rye, fine	Flour, rye
Corn meal, kiln dried	Corn meal, northern	Corn meal, kiln dried
Barley	Wheat, Genesee	
Rye	Rye, northern	
Oats	Oats, northern	
Corn, Indian, northern	Corn, northern	
Bread, ship		Ship biscuits
		Boston crackers (2 kinds)
Rice	Rice, ordinary	Rice
Beans, white		
Starch, American		
	Cod, dry	
Mackerel, bay	Mackerel, No. 1.	Mackerel (3 kinds).
Pork, Cargo No. 1.	Pork, mess	Pork, salt mess
	Pork, prime	
Beef, mess	Beef, mess	Beef, salt mess
	Beef, prime	
Bacon		Bacon, clear
	Hams, smoked	Ham, sugar-cured
Tallow, American	Tallow, American	Tallow
Lard, hog's	Lard	Lard (2 kinds)
Butter, 1st quality	Butter	Butter
Cheese, American	Cheese	Cheese
Raisins, bloom	Raisins	Raisins
Figs, cask	Figs, Smyrna	
Lemons	Prunes, Bordeaux	
Currants		Currants
Almonds, soft shelled		
Coffee, ordinary	Coffee, Java	Coffee, Rio, fair
Tea, Hyson	Tea, Young Hyson	
Molasses, Havana	Molasses, Havana	Molasses (2 kinds)
Sugar, Havanna, white	Sugar, New Orleans	Sugar (3 kinds)
Sugar, Havanna, brown	Sugar, loaf	
Hops	Hops, first sort	
Salt, Liverpool	Salt, Liverpool, fine	Salt (5 kinds)
Pepper	Pepper	Pepper
Nutmegs	Nutmegs	Nutmegs
Cinnamon		
Cloves		
Ginger, ground		
Pimento		
Nankins		
Cotton, Georgia, upland	Cotton, upland	Cotton, upland
		Broadcloths
		Calico
	Wool, common	
	Wool, Merino	
	Wool, Pulled, No. 1.	
Hides, dried	Hides, La Plata	Hides
Leather, foal, dried	Leather, Hemlock	Leather
	Furs, Beaver, Northern	Sole leather
Copper, in sheets	Copper, sheeting	Copper, sheet
		Copper, ingot
Iron, pig, Phil bar, and Russia	Iron, English bar	Iron, rods
Lead, bar	Lead, pig	Lead, pig (2 kinds)
Steel		
Shot		Drop shot, lead
Nails		Nails, cut
Tin, in sheets		Zinc, sheet
Cordage, American		Rope (2 kinds)
		Wood screws
		Spelter
		Quicksilver

COMPARATIVE TABLE OF COMMODITIES USED—(Concluded).

Boston *Gazette* and Boston *Patriot*. 1801–1825.	Report of Secretary of Treasury. 1825–1840.	Aldrich Report. 1840.
Lumber, boards, clear		Lumber, hemlock
		Boards (4 kinds)
		Shingles (3 kinds)
Wood, mahogany, bay		Flooring
Varnish, white		
Lime		Lime
Glue, American		
Tar		Tar
Turpentine	Turpentine	Turpentine
Pitch		
Rosin	Rosin, common	
	Glass, American	Glass, window
Staves, W. Oak, pipe		
Hoops		
Copperas		Copperas
Brimstone		Brimstone, crude
Alum		Alum, lump, crystal
Gunpowder, American	Gunpowder, American	
Oil, linseed	Oil, whale	Oil, linseed
Pot ashes		Soda ash
Flaxseed		Flaxseed
Salt petre, rough		
		Calomel
		Mercury
		Muriatic acid
		Opium
		Quinine
		Sugar of lead (2 kinds)
		Sulphuric acid
		Blue vitriol
Candles, tall, mo. Am.	Candles, mould	
Coal, pit, American	Coal Schuylkill	
Sheeting, Russia	Sheeting, Russia, brown	
Soap, castile	Soap, castile	Soap, castile
Tobacco, James River	Tobacco, Kentucky	
	Tobacco, mfg. No. 1.	
	Whiskey, domestic	Alcohol
	Brandy, Cognac	
	Rum, Jamaica	
	Gin, Scheidan	
Wine, Madeira	Wine, Madeira	
	Wine, Port	
Whalebone, long	Whalebone, slab	
Corks, Lisbon		
Bees wax		
Hemp		
Twine, sewing		
	Olives	
	Paint, red lead	
	Linseed	
	Indigo, Manillo	
		Matches
		Tubs, wooden
		Pails, wooden
		Carpets
		Tickings
		Carbonate of lead
		Putty
		Doors, pine

It should be noted that the prices used in this study are wholesale prices quoted necessarily in United States money. The retail prices of local mercantile establishments often

varied greatly according to the depreciation of the bank notes presented in payment for the goods. This was particularly true in 1814, when banks generally, except in New England, suspended specie payments. Since the value of the depreciated bank notes ranged all the way from par to almost nothing, it is evident that local retail prices might be almost anything, depending on the currency used. Thus, in the war period, retail prices elsewhere soared much higher than in New England where they remained on a specie basis. But it is evident that this fact would have nothing to do with the real level of prices quoted in United States money.

The question of the causes underlying the price curve presented herewith, is not here entered into. The period is complicated by a number of factors, principally the Berlin and Milan Decrees; the Orders in Council; the Embargo Act; the War of 1812; the great growth of state banks between the termination of the first United States Bank in 1811 and the establishment of the second United States Bank in 1816, tending to raise prices by increasing the amount of currency in circulation; the bank crash in New England in 1809; the crises of 1814, 1819, 1825, and 1837; and the monetary measures of the Jackson and Van Buren administrations.

THE PRESENT POSITION OF INFANT MORTALITY: ITS RECENT DECLINE IN THE UNITED STATES.

By Henry Horace Hibbs, Jr., *Sometime Fellow in Research, Boston School for Social Workers.*

The term infant mortality, according to the generally accepted usage, is employed to designate the deaths of infants under one year of age. The problem is measured by an infant mortality rate which is an expression of the proportion of infants dying under one year of age to 1,000 births or, when birth statistics are not available, to 1,000 population under one year of age. The deaths of infants at this early age is made the object of special study because they constitute such an enormous proportion of the deaths at all ages. In no other period of life do deaths occur with such frequency. This is evident upon examination of any of the bulletins on mortality statistics issued annually by the Bureau of the Census for the registration area of the United States. Thus, Bulletin 109 shows that 27 per cent. of the total number of deaths which occurred in the registration area in 1910 were of children under five years of age and 19 per cent., or almost one fifth, were of infants under one year of age. In 1911, 1912, and 1913 the deaths of infants constituted 18 per cent. of the deaths at all ages.

That the death rate for infants and young children should be greater than for other persons is not, however, surprising. As a recent English writer says, "The young of all animals are more susceptible than the adult to the influence of the environment and the approach of death. Hence, it is inevitable that, even under the most favourable circumstances the deaths of infants will furnish a large contribution to the bills of mortality."* It is not the mere fact of excess but the tremendously greater excess of deaths which occur during the first year of life that constitutes the problem of infant mortality.

Infant Mortality in Foreign Countries. The world-wide significance of the problem will be evident upon examination

* George Newman, M.D.: Infant Mortality. London, 1906, p. v.

of the following figures showing the infant mortality rate per 1,000 births for the foreign countries for which statistics are available during the five year period from 1906 to 1910:*

Chile	315	Finland	117
Hungary	204	Switzerland	115
Jamaica	191	The Netherlands	114
Ceylon	189	Scotland	112
Prussia	168	Denmark	108
Servia	154	Ireland	94
Italy	153	Sweden	78
Belgium	141	Australian Commonwealth	78
Ontario	127	Norway	70
France	126	New Zealand	70
England and Wales	117		

Thus, in one third of the countries of the world for which statistics are available the infant mortality rate was over 150, while in about one half it was over 125. In only 5 of the countries, 3 in Europe and 2 in Australasia, was the rate less than 100 deaths per 1,000 births. Expressed in another way this means that, out of every 1,000 children born in countries like Hungary, Prussia, and Italy, from 150 to 200 die before reaching the end of the first year of life; out of every 1,000 born in countries like England, Scotland, Switzerland, and The Netherlands, from 110 to 120 die before reaching this age; while in countries like Ireland, Sweden, Norway, and Australia, from 70 to 100 die before they are a year old.

Infant Mortality in the United States. Figures comparable with these exist for only a few of the states and cities of the United States. In 1911 the Bureau of the Census reported that only in the cities of Washington and New York and in the states of Pennsylvania and Michigan and the six New England States could the registration of both births and deaths be regarded as sufficiently complete (amounting to at least 90 per cent. of the total) to make possible the calculation of an accurate rate of infant mortality based on the ratio of births to deaths. This area comprises the provisional "registration

*Seventy Third and Seventy Fourth Annual Reports of the Registrar General for Births, Deaths, and Marriages in England and Wales (pp. xciv and 103–15 respectively).

area" of the United States for births and deaths. The infant mortality rates per 1,000 births for these states and cities in 1910 were as follows:

Rhode Island.....	158	Connecticut.....	127
New Hampshire...	146	Michigan.....	124
Pennsylvania.....	140		
Maine.....	135	Washington, D. C.	152
Massachusetts....	131	New York, N. Y..	125

Besides these states and cities where the registration of both births and deaths are regarded by the census office as being sufficiently or "fairly complete," there are others where the registration of deaths only is regarded as sufficiently complete to be included in what is known as the registration area for deaths. In 1910, 22 entire states and a large number of cities in non-registration states were included in this area. Since complete birth statistics are not available, it is not possible to calculate an infant mortality rate for this area in the ordinary manner—by computing the ratio of deaths to 1,000 births. To overcome this difficulty, the Bureau of the Census in its report on mortality statistics for 1910 employed an infant death rate based on the proportion of deaths to 1,000 population under one year of age in 1910. This method is confessedly inaccurate because the enumeration of the population under one year of age is never complete and entirely accurate. Yet it is the best method available for studying the distribution of infant mortality in the United States and, although its crudities should serve as a caution against drawing too fine conclusions from its use, its defects should not be over-emphasized. The following table shows the death rate per 1,000 population under one year of age for the registration states:*

Utah	82.3	Ohio.....	115.9
Washington	84.3	Michigan	127.5
Kentucky.....	87.9	Maine	140.4
Montana.....	90.4	New York.....	143.6
California.....	92.2	Connecticut ...	143.7
Minnesota.....	92.4	New Jersey....	148.8

* Bureau of the Census: Bulletin 112, p. 24.

Missouri.....	*96.7	Pennsylvania..	149.7
Colorado.....	104.5	Maryland.....	152.1
Indiana.....	106.9	Massachusetts.	160.8
Wisconsin.....	108.0	New Hampshire	164.9
Vermont.....	109.4	Rhode Island..	181.5

Thus the infant death rate per 1,000 population varied from less than 85 in the Western states, Utah and Washington, to 165 and 182 in the two New England states, Rhode Island and New Hampshire. It was 127.6 for the entire group of registration states considered as a whole.

It would be of interest to compare the rates for this group of states with those for foreign countries given in a previous table. But this is, of course, impossible since infant death rates based on population can not be compared with the true infant mortality rate based on births. In the 1911 bulletin on "Mortality Statistics," the Bureau of the Census estimated on the basis of the figures quoted here and others, that the infant mortality rate per 1,000 births for the United States as a whole was about 124. Comparing this estimate with the computed rates for the foreign countries given in the preceding table, it will be seen that the rate of infant mortality in the United States is lower than in such countries as Chile, Hungary, Jamaica, Prussia, Servia, and Italy; about equal to the rate for the province of Ontario and for France; higher than the rate for England and Wales, Scotland, Finland, Switzerland, and Denmark; and considerably higher than the rate for Ireland, Sweden, Norway, New Zealand, and the Australian Commonwealth.

The figures quoted in the previous table also are of value in that they show the relative position of infant mortality in the different states and sections of the United States. The states included in the registration area are arranged in an ascending order according to their infant death rates. An examination of this table at once reveals the fact that in general the lowest rates are to be found in the Western and the highest in the Eastern states, with the rates for the North-Central (or Middle-Western) states in between. Thus, the

* Figures for deaths for 1911, first year of operation of state law.

average infant death rate for the 5 Western and Mountain states included was 91, for the 6 North-Central states, 108, and for the 9 New England and Middle-Atlantic states, 149.*

The same conclusion that in 1910 infant death rates were lowest in the Western part of the registration area and highest in the Eastern, with the Middle-West in between, is also to be drawn from an examination of the following table showing the infant death rate per 1,000 population under one year of age for the larger cities of the registration area:†

Oakland, Cal.....	94.8	Dayton, Ohio.....	146.8
Seattle, Wash.	100.4	Cleveland, Ohio...	147.2
Portland, Ore.	105.3	Cincinnati, Ohio ..	149.8
Los Angeles, Cal...	110.7	Jersey City, N. J..	153.2
San Francisco, Cal.	113.6	New Orleans, La. .	154.9
Toledo, Ohio.....	125.0	Atlanta, Ga.....	155.3
Cambridge, Mass..	126.1	Bridgeport, Conn..	155.5
St. Paul, Minn. ...	130.8	Philadelphia, Pa...	162.2
Birmingham, Ala..	133.0	Albany, N. Y.....	162.9
Louisville, Ky.....	134.0	Boston, Mass.....	165.5
Denver, Col.....	134.7	Worcester, Mass...	168.0
Grand Rapids, Mich.	134.8	Kansas City, Mo..	170.4
New Haven, Conn..	134.9	Milwaukee, Wis. ..	172.0
Nashville, Tenn....	135.1	Providence, R. I...	173.7
St. Louis, Mo.....	135.8	Syracuse, N. Y.	176.4
Chicago, Ill.	139.5	Pittsburg, Pa.	179.6
Omaha, Neb.....	140.0	Buffalo, N. Y.....	180.9
Columbus, Ohio...	140.4	Washington, D. C.	194.6
Spokane, Wash.....	142.4	Detroit, Mich.....	204.8
Indianapolis, Ind...	144.8	Baltimore, Md....	209.6
Newark, N. J.	145.8	Richmond, Va...	229.3
New York, N. Y...	146.2	Fall River, Mass..	259.5
Paterson, N. J.....	146.7	Lowell, Mass.	261.0

* As only 2 of the 16 Southern states were included in the registration area in 1910, no comparison of the incidence of infant mortality in this with other sections of the country is possible.

† Bureau of the Census: Bulletin 112— Mortality Statistics, p. 24.

The average infant death rate in 1910 per 1,000 population under one year of age was 115 for the 7 Western and Mountain cities included in the table, 149 for the 14 North-Central (or Middle-Western) cities, 165 for the 7 Southern cities, 162 for the 9 Middle-Atlantic cities, and 181 for the 8 New England cities.

This study of the position of infant mortality in the United States and foreign countries shows the seriousness and world-wide significance of the problem. It also shows how the infant mortality rate varies in different parts of the civilized world. Thus, the rate has been found to be much lower in Australia than in Europe. Among the European countries it was lowest in Norway, Sweden, Ireland, and Denmark and highest in Russia, Prussia, Hungary, and Italy. Turning to a single country, the United States, and substituting the use of the infant death rate per 1,000 population under one year of age for that of the infant mortality rate per 1,000 births, the same wide variation was revealed, the ratio of infant deaths to population being considerably less in the Western than in the Eastern parts of the registration area. Further examination of the tables also showed that the ratio varies just as widely when the cities of any state or country are compared. The examination of the report of the health department of almost any city that requires the registration of births and deaths will reveal the same variation by wards—and even by blocks, if figures are given for such small areas.

This wide variation in rates of infant mortality for different countries, states, and cities constitutes a fact of fundamental importance in the study of the subject. Out of it arise questions that at once bring us face to face with the relationship between social and industrial conditions and infant mortality. Why this wide variation in the geographic distribution of infant deaths? Why is the infant death rate lower in one country than another, in certain cities of the same country than others, in certain wards of one city than in others? Why, indeed, should the death rate for little children in the first year of life so far exceed the rate for older children and adults? All of these questions require for their answer some knowledge of the causes of infant mortality and their rela-

tion to industrial, domestic, and social conditions. But with this aspect of the problem this paper cannot deal.*

The Recent Decline of Infant Mortality in Foreign Countries. Since 1881, the first year for which statistics are available for most countries, there has been a noticeable decline in infant mortality in most foreign countries and cities and, since 1900, in most of the states and large cities included in the registration area of the United States. The following table shows this decline for the principal foreign countries for which statistics are available:

PER CENT. OF DECREASE IN THE INFANT MORTALITY RATE PER 1,000 BIRTHS FOR THE PRINCIPAL FOREIGN COUNTRIES FOR WHICH STATISTICS ARE AVAILABLE BETWEEN 1881-85 AND 1906-10. (a)

Country.	1881-1885.	1906-1910.	Per Cent. of Decrease.
EUROPE			
Hungary	250 (b)	204	18.4
Prussia	207	168	18.8
Italy	185 (b)	153	17.3
Servia	157	154	1.9
Belgium	156	141	9.6
France	167	126	24.6
England and Wales	139	117	15.8
The Netherlands	181	114	31.5
Switzerland	171	115	32.7
Finland	162	117	27.7
Sweden	116	78	32.8
Scotland	117	112	4.3
Denmark	135	108	20.0
Ireland	94	94	0.0
Norway	99	70	29.3
AUSTRALASIA:			
The Commonwealth	125	78	37.6
New Zealand	90	70	22.2

(a) Seventy Third and Seventy Fourth Annual Reports of the Registrar General for Births, Deaths, and Marriages in England and Wales (p. xciv and pp. 105-15 respectively).
(b) Figures for 1881-85 not available: those given are for 1891-95.

The rate of infant mortality for every country included in the table declined during this period of 30 years with the single exception of Ireland where, although the rate for both periods remained the same, it was at a very low point—94 deaths

* In other recent articles the writer has discussed this question of the relation of social conditions to infant mortality. See "Infant Mortality and the Size of the Family," QUARTERLY PUBLICATIONS OF THE AMERICAN STATISTICAL ASSOCIATION, September, 1915; "Infant Mortality and Urban, Housing, and Living Conditions," Journal of Sociologic Medicine, October, 1915; "The Relation of Economic and Industrial Conditions to Infant Mortality," Quarterly Journal of Economics, November, 1915; and "The Influence of Prenatal Conditions on Infant Mortality," Proceedings of the Southern Sociological Congress, 1915.

per 1,000 births. The most notable decrease was in New Zealand and Australia. In the former the rate fell from the already low point of 90 deaths per 1,000 births to 70—a decrease of 22.2 per cent.—and in the latter from 125 to 78—a decrease of 37.6 per cent. The decline was also notable in Norway, Sweden, and Denmark and to a lesser extent in England and Wales. The absolute decrease was also great in Switzerland, the Netherlands, France, and Finland, but the rate was very high for each of these countries at the beginning of the period.*

The Registrar-general for England and Wales, from whose annual reports the preceding table was compiled, also gives figures showing the decline in the rate of infant mortality in the principal foreign cities since 1881–85. Space will not permit quotation of these in detail but the fact should be noted that in each of the cities included, with one exception (Trieste, Hungary), the rate of infant mortality declined during the period under consideration. The most notable decrease was in the three Dutch cities, Amsterdam, The Hague, and Rotterdam; the two Australian cities, Sydney and Melbourne; and the cities of Norway and Sweden, Stockholm and Christiania, in each of which the rate fell to a point below 100 deaths per 1,000 births—a record, as shown in the preceding table, also attained by each of the countries in which these cities are situated. The absolute decrease was also great in the two Prussian cities, Munich and Berlin, and the Hungarian city of Budapest, but the rate for each of these cities was very high at both the beginning and the end of the period.

The Decline in Infant Mortality in the United States. Unfortunately no series of infant mortality rates at all comparable with those just shown for foreign countries can be presented

* In view of the fact that, as has been frequently pointed out, the *apparent* decline in the rate of infant mortality in any country in a period of years may be affected by the increase in the per cent. of births which are registered, the figures given in the table may not in all cases be strictly comparable. For instance, if in three countries, in each of which the proportion of births registered in 1881 was 90 per cent., it should happen that the proportion registered should gradually increase in each *but unequally* so that in the first 92 per cent. of the births which occurred in 1910 were registered, and in the second 95 per cent., and in the third 99 per cent., the decline in the rate of infant mortality between these two years would not be strictly comparable unless the factor of varying perfection in birth registration were allowed for. This difficulty probably is not of sufficient importance to require its consideration here even if sufficient material bearing on the comparative efficiency of birth registration in foreign countries in the last thirty years were available.

for the United States. Figures are available, however, for Massachusetts and Boston for the same period, 1881–85, to 1906–10 and for three later years, 1911–13. Also, the per cent. of decrease in the infant death rate per 1,000 population under one year of age between 1900 and 1911 has been calculated by the Bureau of the Census for the registration area and the larger registration cities. The following table shows the decline in the infant mortality rate for Massachusetts and Boston since 1881 and the per cent. of decrease in the rate between 1881–85 and 1909–13:

PER CENT. OF DECREASE IN THE INFANT MORTALITY RATE PER 1,000 BIRTHS FOR THE COMMONWEALTH OF MASSACHUSETTS AND THE CITY OF BOSTON BETWEEN 1881-85 AND 1909-13. (a)

Years.	Massachusetts.	Boston.
1881–85	160	186
1886–90	161	178
1891–95	161	167
1896–1900	153	151
1901–05	138	138
1906–10	133	133
1909–13 (b)	121	120
Per cent. of decrease	24.4	35.5

(a) Compiled from the Massachusetts annual reports on births, deaths, and marriages and the annual reports of the Health Department of Boston.
(b) Figures for five year period are not available.

It will be noted that the infant mortality rate in this period of thirty-three years decreased over 24 per cent. in Massachusetts and about 36 per cent. in Boston. It will also be noted that the decrease was especially marked during the past few years.

The nearest approach to an accurate determination of the position of infant mortality in the other states and cities of the United States is to be found in a table recently presented by the Bureau of the Census, and herewith reproduced in part, which shows the per cent. of decrease in the infant death rate per 1,000 population under one year of age between the census year 1900 and the calendar year 1911 for the states and large cities of the registration area. It should be noted, however, that the rates given in this table are infant death rates calculated upon the basis of infant deaths to 1,000 population under one year of age and not according to the usual method of the ratio of deaths to 1,000 births.

PER CENT. OF DECREASE IN THE INFANT DEATH RATE PER 1,000 POPULATION UNDER 1 YEAR OF AGE BETWEEN THE CENSUS YEAR 1900 AND THE CALENDAR YEAR 1911 FOR THE STATES INCLUDED IN THE REGISTRATION AREA IN 1900 AND FOR CITIES OF 400,000 POPULATION OR OVER. (a)

Area.	Census Year: 1900.	Calendar Year: 1911.	Per Cent. of Decrease.
States included in the registration area in 1900 (b)	159.3	129.5	19
Rhode Island	197.9	138.6	30
Massachusetts	177.8	143 3	19
New Hampshire	172.0	150.3	13
New Jersey	167.4	131.5	21
New York	159.8	128.8	19
Connecticut	156.8	130.9	17
Maine	144.1	110.9	23
Vermont	122.1	102.0	16
Michigan	121.3	111.4	8
Cities of 400,000 population or over in 1910. (c)			
Baltimore	235.1	189.2	20
Philadelphia	201.9	141.9	30
Detroit	201.2	168.8	16
Boston	194.1	160.9	17
New York	189.4	130.6	31
Cleveland	185.5	123.7	33
Pittsburg	179.8	141.4	21
St. Louis	162.4	123.8	24
San Francisco	152.2	104.8	31
Buffalo	150.9	140.6	7
Chicago	146.6	123.3	16

(a) Twelfth Annual Report of the Bureau of the Census on Mortality Statistics for the year 1911, p. 24.
(b) Includes District of Columbia.
(c) Space does not permit the quoting of rates for smaller cities.

From this table it will be noted that in this period of 11 years the ratio of infant deaths to 1,000 population under one year of age decreased nearly one fifth (19 per cent.) in this group of registration states. The largest decrease shown in the rate for any of the states was in that for Rhode Island (30 per cent.) and the least in that for Michigan (8 per cent.). In all the cities included in the table the infant death rate also showed a decline—ranging from 33 per cent. in Cleveland to 7 per cent. in Buffalo. The fact that this comparison relates to only two individual years and that complete returns of deaths of infants under one year of age may not always have been made, coupled with the fact that the number of infant deaths per 1,000 population under one year of age does not furnish as satisfactory a basis for the study of infant mortality as the number of such deaths per 1,000 births, tends to diminish somewhat the value of the figures given in

the table. Yet, in spite of these limitations, these figures, taken in conjunction with those previously given for Massachusetts and Boston, show that in all probability there has been a marked reduction in infant mortality in this country in recent years.*

The Decline in the Infant Mortality Rate Compared with that in the Gereral Death Rate for All Ages. The extent of the decline in the mortality rate for infants under one year of age can not be fully appreciated until it is compared with the decline in the death rate for other age periods. The following table compares the decline in the infant mortality rate shown in preceding tables with the decline in the general death rate for all ages during the same periods:

PER CENT. OF DECREASE IN THE INFANT MORTALITY RATE PER 1,000 BIRTHS AND IN THE GENERAL DEATH RATE FOR OF ALL AGES PER 1,000 POPULATION BETWEEN 1881-85 AND 1891-95 AND BETWEEN 1896-1900 AND 1906-10, FOR THE PRINCIPAL FOREIGN COUNTRIES. (a)

Country.	Per Cent. of Decrease Between 1881-85 and 1891-95		Per Cent. of Decrease Between 1896-1900 and 1906-10	
	General Death Rate.	Infant Mortality Rate.	General Death Rate.	Infant Mortality Rate.
Hungary	3.9	—	10.4	6.8
Prussia	10.2	1.0	17.6	16.4
Italy	6.6	—	8.3	8.9
Servia	+15.2	+8.7	+35.6	3.1
Belgium	2.4	+4.9	+26.7	10.8
France	+0.4	+2.3	7.2	20.8
England and Wales	3.6	+7.9	16.9	25.0
The Netherlands	8.4	8.8	16.8	24.5
Switzerland	7.0	9.3	+30.4	19.6
Finland	7.7	10.5	8.4	15.8
Sweden	5.1	11.2	11.2	22.8
Scotland	3.1	+7.1	10.6	13.2
Denmark	+1.1	+2.2	16.5	18.2
Ireland	+2.7	+7.8	4.4	11.3
Norway	2.3	1.0	11.5	27.1
New Zealand	7.3	3.3	+1.0	12.5
The Australian Commonwealth	15.3	12.8	15.7	30.4

A plus sign (+) denotes an increase.

(a) Compiled from the Seventy Third and Seventy Fourth Annual Reports of the Registrar General for Births, Deaths, and Marriages in England and Wales for the years 1910 and 1911. The Twelfth Annual Report of the Bureau of the Census on Mortality Statistics for 1911 quotes in detail the general death rates from which the per cents. of decrease in this table were compiled.

* This is the conclusion arrived at in the Twelfth Annual Report of the Bureau of the Census on Mortality Statistics for the year 1911, p. 24, and expressed in the following words: "There has been a marked reduction in the infant death rate in recent years."

The first thing to be noted upon examination of the above table is the much greater per cent. of decrease in both the general death rate and the infant mortality rate in the last than in the first half of this period of thirty years. Thus, from 1881–85 to 1891–95 the infant mortality rate decreased in only 8 of the 15 countries for which rates could be obtained, while from 1896–1900 to 1906–10 it declined in every country included in the table. Moreover, the average per cent. of decrease for all countries in the latter period was twice as great as in the former. The same variation is also shown in the decline of the general death rate in the two periods but to a somewhat lesser extent.

By comparing the extent of the decline in the infant mortality rate with that in the general death rate it will be seen that in the first half of the period the greater decline occurred in the general death rate, while during the second half the greater decline occurred in the infant mortality rate. Thus, from 1881–85 to 1891–95, a greater per cent. of decrease in the infant mortality rate occurred in only 4 of the countries included in the table, while from 1896–1900 to 1906–10 a greater per cent. of decrease failed to occur in only 2 countries.

A similar comparison extending over the same periods can be made for Massachusetts, and it shows the same results. Thus, from 1881–85 to 1891–95, the general death rate for Massachusetts decreased 0.5 per cent. and the infant mortality rate increased 0.6 per cent., while during the period from 1896–1900 to 1906–10 the former rate decreased 10.5 per cent. and the latter 13.1 per cent.

That this greater decline in the infant mortality rate than in the general death rate during recent years is probably typical for this country is shown in the following table, which compares the per cent. of decrease between 1900 and 1911 in the general death rate and the infant death rate for the states included in the registration area in 1900:

PER CENT. OF DECREASE IN THE INFANT DEATH RATE PER 1,000 POPULATION UNDER 1 YEAR OF AGE AND THE GENERAL DEATH RATE FOR ALL AGES PER 1,000 POPULATION BETWEEN 1900 AND 1911, FOR THE STATES INCLUDED IN THE REGISTRATION AREA IN 1900. (a)

State.	General Death Rate.	Infant Mortality Rate.
All States (b)	14	19
Rhode Island	25	30
Maine	13	23
New Jersey	17	21
New York	14	19
Massachusetts	17	19
Connecticut	15	17
Vermont	9	16
New Hampshire	13	13
Michigan	11	8

(a) Twelfth Annual Report of the Bureau of the Census on Mortality Statistics for the year 1911, pp. 22 and 25. The general death rates are "corrected on the basis of the standard million of England and Wales."

(b) District of Columbia included in both rates and Indiana in the general death rate in addition to the states mentioned.

During this period of 11 years a greater decline in the infant death rate than the general death rate for all ages occurred in all of the 9 states included in the above table except 2, New Hampshire and Michigan. In the former the decline in the two rates was exactly equal. All evidence seems to point, therefore, to the conclusion that the decline in the infant mortality rate during the last 10 or 15 years has been greater than that in the general death rate for all ages.

Before leaving this subject it will be advisable to compare the decline in mortality by age. This is possible from the figures given in the following table comparing the per cent. of decrease in the death rate for persons of different ages between 1900 and 1911 for the group of registration states as constituted in 1900:

PER CENT. OF DECREASE IN THE DEATH RATE PER 1,000 POPULATION FOR CERTAIN AGE GROUPS BETWEEN 1900 AND 1911, FOR THE STATES INCLUDED IN THE REGISTRATION AREA IN 1900. (a)

All ages..... 13

Under 1 year	22	25 to 34 years	23
1 to 4 years	35	35 to 44 years	9
5 to 9 years	32	45 to 54 years	3
10 to 14 years	27	55 to 64 years	+4
15 to 19 years	27	65 to 74 years	+3
20 to 24 years	26	75 years and over	0

A plus sign (+) denotes an increase.

(a) Twelfth Annual Report of the Bureau of the Census on Mortality Statistics, p. 22.

An examination of these figures shows that the death rate for all age groups under 55 decreased between 1900 and 1911. The greatest decrease was for the age group 1 to 4 years, the per cent. of decrease falling off with each succeeding age group until the period from 55 to 64 years was reached, this and the next group showing a small increase. The death rate above 75 years was practically the same in each period. The per cent. of decrease in the mortality of the first year of life was noteworthy (22) but it was exceeded by that of the years of both childhood and adolescence, being about equal to that for the age group 25 to 34 years and greater than that for all succeeding groups.

It has thus been shown that since 1881 the rate of infant mortality has been declining in practically all European countries for which statistics are available, in the Australian Commonwealth and New Zealand, and in Massachusetts and Boston in this country. This decline in infant mortality has been especially marked in the last twelve or fifteen years during which period it has practically everywhere exceeded the decline in the general death rate for all ages. It has also been shown that between 1900 and 1911 a marked decline in the infant death rate per 1,000 population under one year of age occurred in the states and large cities of the registration area. Here, too, the decline in the infant death rate was with one or two exceptions found to be greater than that in the general death rate. On comparing the per cent. of decrease in the death rate for infants under one year of age with that for other ages it was found that the decline in infant mortality was less than that for children or young persons under 25 years of age but about equal to that for persons in the age group 25 to 35 years and greater than that for persons over 35. It appears, therefore, that the decline in the infant death rate has in general been greater than the decline in the adult death rate; but the decline in the infant death rate has not been as great as the decline in the death rate for persons in the years of childhood and adolescence.

REVIEWS AND NOTES.

NOTE.

A Division of Statistical Research was organized in the New York City Health Department within the Bureau of Records on June 1, 1915 The creation of such a division has long been contemplated by the Department and repeatedly recommended by special committees.

It is hoped that the new division will justify its existence from the start and thus secure a reasonable amount of support in the next budget. The new division hopes to take over the statistical work of the different bureaus of the department and in this way to secure a saving of labor that can be devoted to the development of the department's statistics. The division will then be able to supply the bureaus of the department with scientific data to guide them in their different fields of health work.

The new division has been organized entirely without additional cost to the city by detailing qualified workers from within the department. Shirley W. Wynne, M.D., Assistant Registrar, Manhattan, has been designated Chief of the new division.

L. I. D.

STATISTICS OF SUICIDE IN SPAIN AND SAXONY.

Under the caption, "*Zur Statistik des Selbstmords,*" the *Deutsches Statistisches Zentralblatt** discusses recent publications on suicide statistics in Spain and in the Kingdom of Saxony.

Spanish suicide statistics, in spite of their detail of presentation, have hitherto obtained but little notice. They were formerly grouped with the general statistics of mortality, but were made the subject of special enumeration by a royal decree of September 8, 1906. As suicide still entails certain legal consequences in Spain, the lower courts were entrusted with the filling out of a questionnaire for every suicide and attempted suicide, and were instructed to send these questionnaires, at the end of every quarter, to the Geographical and Statistical Institute for compilation. These forms do not contain the names of suicides. They record nationality, home, temporary or permanent character of stay in locality, literacy, manner of committing suicide (9 titles), and cause of suicide (13 titles).

The first comprehensive compilation for Spain has appeared under the title, "*Estadistica de Suicidio en España*" (Madrid, 1913) for 1906–1911. It analyzes the material by years, provinces and large cities, age, conjugal condition, time of suicide, etc. Of course the criminal nature of the act in Spain is certain to have made the figures too low. The tables are accompanied by comparisons with other European countries. It appears that, during the period of observation, there were 74 suicides and attempted

* Vol. II, No. 1. Leipsig, January, 1915.

suicides per million population. It is noteworthy that these cases occurred, for the most part, among persons of the higher cultural grades.

Suicides occurring in Saxony since 1830 are discussed by Dr. Georg Radestock in the *Zeitschrift des Königlichen Sächsischen Statistischen Landesamtes*, 1913, No. 2, p. 275 ff. He depicts the evolution of suicide statistics and of the methods of enumerating them. Reliable data are obtained by means of a comparison of police reports with registry office mortality records. Of course a certain percentage of false reports by physicians must be reckoned with.

In the recent figures it may be pointed out that suicide is the second cause of death, with respect to frequency, among men aged 15–29; it is exceeded only by tuberculosis. It is third in order of frequency among men aged 30–39. It is almost twice as frequent at ages 30–59 as at ages 15–29. If these figures are accurate, then the incidence of suicide in Saxony is much higher than it is in the Registration Area of the United States, where suicide ranks seventh in the list of causes of death among males in age period 15–29 and sixth in age period 30–39. In the Industrial mortality experience of the Metropolitan Life Insurance Company, suicide in the first age period ranks eighth and in the second age period ranks fifth.

The author points out that city dwellers usually seek rural surroundings to commit suicide, and discusses the influence of the time of day and of the day of the week; Saxony is perhaps unique in studying the bearing of these factors upon the question. In 1911 most of the suicides occurred on Saturday, in contrast to the predominant occurrence of fatal accidents on Monday. Enumeration by occupation brought interesting results. Analysis by religious confession was incomplete, but steps are being taken to remedy this defect.

O. Kürten has treated the same subject in "*Statistik des Selbstmordes im Königreich Sachsen*," appearing as Supplement No. 3 of the *Deutsches Statistisches Zentralblatt*. He devotes himself rather to the general bearings of the subject than to its many detailed aspects. Among the interesting points, however, which he brings to light, is the fact that suicide attains its greatest frequency, not in large cities, but in towns of 16,000–25,000 inhabitants. He is skeptical with regard to studies of the correlation of suicide with criminality, food-prices, economic and political crises, etc.

LOUIS I. DUBLIN.

Manual for Health Officers. By J. Scott MacNutt. Published by John Wiley and Sons, New York, 1915. Cloth, $3.00.

This volume is a useful addition to the equipment of American public health officials. The author is to be congratulated on a work which so completely puts at the disposal of health workers the best thought and practise in their important field. The point of view is at once comprehensive and radical, as might be expected from Mr. MacNutt whose admin-

istration as health officer of Orange, New Jersey, attracted wide attention. It is the newer public health that is emphasized; the public health which embraces not only the best medical practise but includes the essentials of sanitary engineering and social science as well. The author's association with Professor Sedgwick and his school is a guarantee of an authoritative and progressive treatment of the subject.

Part I contains four useful chapters which outline the historical development of the public health movement in America, local, state, and federal. The general scope of the new public health work is clearly outlined.

Part II, the rest of the volume, gives a detailed treatment of the daily routine of the health officer in his control of the communicable diseases; considers such important matters as child hygiene, the milk and water supplies of communities, housing and industrial hygiene, and the question of nuisances. All of these problems are considered in the light of the most authoritative practise in America and abroad.

To the readers of this QUARTERLY, the most interesting sections of the volume will be those which treat of vital statistics. In general, it may be said that the author appreciates the value of vital statistics as an aid to public health work. He places emphasis in every chapter on the need of good vital statistics as a proper basis for scientific health administration, whether it be in the control of the communicable diseases or of infant mortality or of the housing problem. In addition, an entire chapter, IX, is given to the subject of vital statistics. This is a skillful although a non-technical treatment of a difficult subject. An exposition is given of the underlying principles of statistics, their purposes and uses and the chief sources of error in their interpretation. But more important are the hints, which are in large measure the result of the author's own experience as a health officer, to guide others through the many difficulties involved in the registration of vital statistics, especially of births and deaths. Excellent instructions are given for the preparation of statistical tables and of the annual health report in accordance with the rules of statistical practise approved by the Section on Vital Statistics of the American Public Health Association. The value of the chapter would have been enhanced, however, by the inclusion of the model tables recommended by Committees of the Section.

This book has already found its place in the literature of American public health work and will grow in usefulness as the movement demanding full-time, well-trained, health officers is extended throughout the country.

LOUIS I. DUBLIN.

Elements of Record Keeping for Child Helping Organizations. By Georgia G. Ralph. New York Survey Associates, 1915. Pp. XII+195.

One who has had any connection with an organization engaged in social work knows that success cannot be attained without good case work. Proper investigation must be supplemented by adequate records. The personnel of an office force is constantly changing and memory is faulty.

It is all too often the case that a conscientious worker connected with an organization for years is suddenly removed and unless the records are in good shape the past becomes more or less a myth. The Russell Sage Foundation deserves the thanks of child helping organizations in this country in that it has made possible the preparation and publication of this volume. If the superintendents of children's institutions would read this volume carefully and profit by it, there would be a big change in the character of their annual reports. There is, of course, a limit beyond which it is impossible for an organization to go in multiplying records, but no organization is justified in caring for children unless they have some record showing the antecedents of the children, their physical and mental condition when received, their history while in the institution, and a record of them since they were placed out. Some of the blanks proposed seem unduly elaborate, but when one sits down to decide what question could be omitted without detracting from the value of the records, he is troubled to make a start. Not only are forms suggested for proper record keeping but cards in present use by a large number of different organizations are included in the volume.

For many workers perhaps the two most valuable chapters are those on methods and devices for making records available and on abstracting the material for annual reports. Many social workers, besides those engaged with children, could read these two chapters with profit. A short bibliography is included and the volume is well indexed. One would receive the impression from reading this volume that the author had not only made a careful theoretical study of the field, but had been confronted with some of the problems through actual work.

WILLIAM B. BAILEY.

"*Statistics Relating to Crime in Chicago.*" By Edith Abbott, Ph.D. 88 pp. Reprinted from Report of a City Council on Crime of the City of Chicago. Chicago School of Civics and Philanthropy, 1915.

In this statistical study of crime in Chicago, Dr. Abbott had done a good piece of work. The figures had to be collected from four different sources and when all the available material was collected the data were still lacking for a complete study of the social status of the offenders. The report is divided into two parts: the first covering the statistics of arrest and trial of offenders, and the second their social condition. There are altogether thirty-nine statistical tables besides a considerable number included in the eight appendices. These tables are well planned and give both absolute numbers and percentages in almost every case. In addition to the tables and accompanying text, there are two general summaries giving the main lessons to be gained from a study of the statistical tables. It is apparent that over one half of the persons brought before the courts of Chicago are discharged without conviction. The waste from such a large number of unnecessary arrests is apparent. Only about 3 per cent. of the total number of persons arrested are sentenced to prison. Out of 109,711 persons arrested in 1913, 2,076 were sentenced to the County Jail

or House of Correction and 2,182 were held for the grand jury. Over 80 per cent. of the commitments to the House of Correction were for non-payment of fines. About 60 per cent. of the persons arrested are under 30 years of age. A careful study of the statistics seems to show that the foreign born population is no more criminal than the native born when proper weight is given to the differences in sex and age distribution.

It is to be hoped that this careful study of criminal statistics in Chicago by Dr. Abbott will lead to the publication by that city of a series of annual reports which will render unnecessary another study of this nature.

W. B. B.

Negro Year Book—An Annual Encyclopedia of the Negro, 1914–1915. By Monroe N. Work, 443 pages. 35 cents. Negro Year Book Publishing Company, Tuskegee Institute, Alabama.

The Negro Year Book for the period 1914–1915 fills the need for a comprehensive presentation of facts about the interests and progress of the negro race in America. It is a permanent record of the essential historical facts relating to negro population, health, religious affiliation, education, biography and social organization. The extensive bibliography at the end of the volume should be especially useful. Mr. Work is to be commended for having brought together such a convenient and useful collection of data.

E. W. Kopf.